ORIGINAL ARTICLE

POL-RISK: an algorithm for 10-year fracture risk prediction in postmenopausal women from the RAC-OST-POL study

Wojciech Pluskiewicz[1], Piotr Adamczyk[2], Aleksandra Werner[3], Małgorzata Bach[3], Bogna Drozdzowska[4]

1 Department of Internal Medicine, Diabetology, and Nephrology – Metabolic Bone Diseases Unit, Faculty of Medical Sciences in Zabrze, Medical University of Silesia, Katowice, Poland
2 Department of Pediatrics, Faculty of Medical Sciences in Katowice, Medical University of Silesia, Katowice, Poland
3 Department of Applied Informatics, Silesian University of Technology, Gliwice, Poland
4 Department of Pathomorphology, Faculty of Medical Sciences in Zabrze, Medical University of Silesia, Katowice, Poland

EDITORIAL

by Catalano et al

ABSTRACT

INTRODUCTION   Fragility fractures are the major consequence of osteoporosis. Thus, fracture risk assessment is an essential part of the diagnostic process in osteoporosis.

OBJECTIVES   The aim of the study was to develop an algorithm for fracture risk prediction.

PATIENTS AND METHODS   Bone status was evaluated in a population-based cohort of postmenopausal women at a mean (SD) age of 66.4 (7.8) years. Subsequently, all participants were contacted by phone once a year (for 10 consecutive years) to update their history of fractures. At the end of the 10-year follow-up, the number of the study participants was 640, of whom 129 had a history of 190 osteoporotic fractures recorded during the study period. Statistical analysis included multistep data preprocessing, feature selection, identification of fracture risk factors, and design of the final model. Logistic regression models were fitted and used for the evaluation of variables from determined feature sets, including global fit measures, as well as individual parameters, such as the Wald statistic and P value, odds ratio, and 95% CI.

RESULTS   The 10-year risk for any fracture depended on the age of the patient, the number of recorded fractures after the age of 40 years, femoral neck bone mass, and the occurrence of falls during the previous year. The achieved equation was incorporated into an algorithm, available at the www.fracture-risk.pl webpage.

CONCLUSIONS   A fracture prediction algorithm was developed in a longitudinal study to enable calculation of the 10-year fracture risk. Identification of patients at a high risk of fracture should be followed by implementation of appropriate treatment strategy to reduce the number of future fractures.

INTRODUCTION   Fracture risk is a major concern in patients with osteoporosis. Fragility fractures are a consequence of osteoporosis and one of the most significant risk factors for future fractures. Therefore, management of patients with osteoporosis should include expert recommendations and structured prophylaxis to avoid the first fracture. Due to a prolonged and silent course of the disease, the ability to predict the risk of fracture is very important. The fracture risk is determined by several factors, including bone mineral density (BMD), age, and the so-called clinical risk factors. Several tools have been designed to establish the fracture risk, including FRAX,[1] the Garvan algorithm,[2,3] and QFracture.[4] The Garvan tool allows for estimating the 5-year or 10-year fracture risk, which is assessed separately for the hip and for any osteoporotic fracture. FRAX is a fracture risk assessment tool designed to estimate the individualized 10-year probability of a hip and any major (hip, spine, arm, and forearm) osteoporotic fracture.

WHAT'S NEW?

Based on the results of the study it was possible to develop an entirely new, original diagnostic tool for predicting the risk of osteoporotic fractures. Owing to a 10-year prospective observation model in an epidemiological, population-representative study cohort, medical practitioners gained the possibility of predicting fractures, which may be useful in everyday clinical practice. Inclusion of the new risk calculator in the diagnostic approach should facilitate the implementation of a therapy adequate to the calculated risk of fractures. Widespread use of the new, easily available diagnostic tool may result in the reduction of the number of new fractures and lower the costs associated with osteoporosis for the health care system.

In 2010, an epidemiological, population-based sample of women was recruited to participate in the RAC-OST-POL study (the name of the study consists of abbreviations from the words RACibórz [a town in Poland], POLand, and OSTeoprosis). The basic epidemiological features of this population sample were presented previously.[5] Using longitudinal data collected during a 5-year follow-up, a new algorithm was introduced in 2017 to determine the 5-year fracture risk in Polish postmenopausal women over the age of 55 years.[6] This algorithm has since been available at the www.fracture-risk.pl webpage, and is known as the POL-RISK calculator.[7] Subsequently, the follow-up period of the study was extended to 10 years.

This paper presents the results of the 10-year longitudinal follow-up. The main aim of the study was to develop a new fracture prediction algorithm.

PATIENTS AND METHODS An epidemiological, population-based sample of postmenopausal women aged over 55 years was analyzed in the follow-up study. At the baseline (May 2010), 625 women at a mean (SD) age of 66.4 (7.8) years were invited to participate. Various factors with a potential influence on the metabolism and bone status were recorded using a specially designed questionnaire, and more than 200 variables were gathered. Each of the 625 participants was randomly selected based on local demographics. Additionally, 353 volunteers submitted a request to participate in the project. The crucial characteristics, such as the mean age, place of residence, employment status, marital status, and education level did not differ between the random sample and the group of volunteers; therefore, many further analyses based on which the 5-year algorithm had been developed were carried out using joint data from both groups. Thus, a total of 978 women eventually formed the baseline group. Part of the baseline assessment of the study participants was bone densitometry, performed with a Lunar DPX device (GE Healthcare, Madison, Wisconsin, United States). The femoral neck (FN) was evaluated. All the measurements were performed by a single operator. The percentage of coefficient variation for FN measurement was 1.6%.[5]

Subsequently, that is, from 2011 to 2020, a single author (WP) contacted each participant by phone once a year (in May) to enquire about any recent fractures with a nontraumatic cause. The patients were asked to report the fracture(s) as well as any fall incidents and the received therapy. Of note, only the events confirmed by a given patient's general practitioner were taken into account. Ultimately, 640 women completed the follow-up. The total dropout rate was 34.5%, while some participants were lost to follow-up due to loss of contact (25.6%), death (7.5%), refusal to cooperate (1%), and the lack of baseline bone density scans (0.4%). To assess whether the dropout effect could significantly bias the process of creating an algorithm for fracture risk assessment, the most important clinical characteristics of the women who did not complete the study were additionally analyzed and compared with the final study group. The women who were lost to follow-up were older in comparison with those included in the final analysis (mean [SD], 67.4 [8.6] vs 65 [6.95] years, respectively; $P <0.001$). However, the groups did not differ significantly with respect to the crucial risk factors for fracture, for example, mean (SD) FN T-score (−1.34 [0.9] vs −1.24 [0.9]; $P = 0.08$), the incidence of fractures before enrolment into the study (27.2% vs 30.4%; $P = 0.3$), the incidence of falls during the year preceding the enrolment (35.7% vs 32.4%; $P = 0.3$), and the frequency of steroid use (7.1% vs 5%; $P = 0.18$).

As many as 190 osteoporotic fractures were identified in 129 women during the follow-up, including fractures of the forearm (n = 81), spine (n = 30), ankle (n = 25), hip (n = 15), arm (n = 13), rib (n = 9), foot (n = 7), clavicle (n = 7), and pelvis (n = 3). During the initial phone interviews conducted in 2011, the participants reported 25 recent fractures, while 21, 18, 16, 13, 26, 14, 26, 16, and 15 fractures were recorded year by year, respectively, in the consecutive years. Single fractures were reported by 91 women, whereas 24 women reported 2, 7 women reported 3, 5 women reported 4, and 2 women reported 5 fractures.

TABLE 1 presents the clinical characteristics of the follow-up population, both for the entire group and for subgroups divided according to the occurrence of fractures. **TABLE 2** presents the prevalence of clinical risk factors, both in the entire group and in the subgroups. The approval for the follow-up project was provided by the Ethics Committee of the Medical University of Silesia in Katowice, Poland (KNW/0022/KB1/9/I/10).

Statistical analysis The present study was conducted to develop a mathematical model for the estimation of low-energy fracture probability in a period of 10 years in postmenopausal women. For that purpose, it was necessary to define a set of model predictors. The steps leading to the development of the final regression model are presented in Supplementary material, *Figure S1*.

TABLE 1 Clinical characteristics of the whole study group and subgroups divided according to the occurrence of fractures

Variable	Whole group (n = 640)	Women affected by fractures (n = 129)	Women unaffected by fractures (n = 511)
Age, y	65 (6.95)	66.61 (7.22)[a]	64.64 (6.82)
Height, cm	156.6 (5.59)	156.8 (5.29)	156.68 (5.66)
Weight, kg	74.9 (14)	73.86 (12.83)	74.65 (14.28)
BMI, kg/m²	30.36 (1.65)	30.1 (5.24)	30.42 (5.48)
Age at menopause, y	49.2 (4.93)	49.27 (4.86)	49.16 (4.85)

Data are presented as mean (SD).

a   Significantly higher than in the women unaffected by fractures (*P* <0.01); other variables did not differ.

Abbreviations: BMI, body mass index

TABLE 2 Prevalence of clinical risk factors in the whole study group and subgroups divided according to the occurrence of fractures (in percentage)

Risk factors, n	Whole group (n = 640)	Women affected by fractures (n = 129)	Women unaffected by fractures (n = 511)
0	27.4	22.5	28.6
1	34.5	30.2	35.4
2	25.9	29.5	25
3	9.1	10.1	8.8
4	2.5	5.4	1.8
5	0.6	1.5	0.4
6	0.15	0.8	–

To ensure the validity and quality of the gathered data, and thus to be able to use machine learning (ML) methods, meticulous data preprocessing was carried out, preceded by anonymization of the data by physicians (ie, removal of sensitive personal information).

Various tools and environments were used for data preprocessing and their further analyses. In particular, PQStat v. 1.8.2.226 (PQStat Software, Poznań, Poland; https://pqstat.pl/), Statistica 13 (StatSoft Inc., Tulsa, Oklahoma, United States; www.statsoft.com), and the R programming environment v. 4.0.3 (The R Foundation for Statistical Computing, Vienna, Austria; https://www.r-project.org/) were used to enable extraction of meaningful information from the collected data. It referred to the technique of data preparation, that is, cleaning, feature selection, and organization of raw data to make them suitable for building and training more accurate predictive models.[8,9]

At that stage, the variables that failed to contain significant and useful information were removed, including, among others, those with many missing records or those of little variability (ie, similar for all or almost all participants) and, eventually, those that had been copies of other variables.

Initially, the medical histories of the 640 patients were described by means of more than 200 features. However, after all the preprocessing steps, the list of data features, analyzed in terms of their impact on the occurrence of fractures, was reduced to 19, including age, height, height loss, body mass index, age of menarche, duration of menses (in years) and of lactation (in months), number of deliveries, presence of comorbidities (such as rheumatoid arthritis, asthma, depression), steroid administration, supplementation of vitamin D and calcium, parental history of hip fractures, BMD of FN determined by dual-energy X-ray absorptiometry and reported in T-score, number of fractures recorded after the age of 40 years, as well as occurrence of falls from standing height and their number over the last 12 months.

Each data record (representing 1 patient) had a label stating whether the patient had or had not experienced low-trauma fractures, which was a predicted (also called dependent, target, or outcome) variable.

To investigate the collinearity between the pairs of variables that remained after the preprocessing phase, a correlation matrix was defined. Its graphical representation, a so-called heatmap, is shown in **FIGURE 1**. The correlation analysis was performed, using 1) Spearman correlation coefficient between 2 continuous variables, 2) Cramer V coefficient between 2 discrete variables, and 3) point biserial coefficient of correlation between continuous and discrete variables.

Additionally, considering the possibility that collinearity may have occurred not only between pairs of variables but also among 3 or even more variables (multicollinearity), the variance inflation factor (VIF) was calculated.[10] According to the recommendations,[11] VIF values greater than 5 indicate a strong correlation between a given variable and other predictor variables; therefore, only 1 of the collinear variables should be left in the model. However, other sources[12,13] provide evidence that collinearity is possible even when VIF is below 5. The former guidelines were considered in the preliminary analyses, however, it was decided to include neither vitamin D nor Ca supplementation in the final model as the VIF values for these parameters were 4.529 and 4.61, respectively.

Simple (univariable) logistic regression was used as a "first-line" model. It was examined whether the model that contained a specific predictor variable was significantly better than the model with the intercept only. In such a way, a preliminary identification of the potential risk factors for a fracture was made. However, when only the variables with a significant association in the univariable analysis are included in a multivariable logistic regression model, important covariates may be disregarded, which could lead to biased estimates and wrong conclusions. Therefore, we constructed a multivariable regression model using forward and backward stepwise regression, both with the following variable feature selection methods: conditional parameter estimates, the probability of the likelihood-ratio statistic, and the probability of the Wald statistic, which gave a total of 2 × 3 = 6 models.

	Age	Height	Height loss	Body mass index	Age at menarche	Menses duration	Number of labors	Lactation duration	Hip fracture in parents	Number of prior fractures	Rheumatoid arthritis	Steroid use	Asthma	Depression	Calcium supplementation	Vit D supplementation	Femoral neck T-score	Number of prior falls	Prior falls
Age	1	−0.148	0.426	0.144	0.137	0.129	0.115	0.135	−0.062	0.034	0.072	0.064	0.024	0.033	−0.018	0.076	−0.155	0.179	0.185
Height	−0.148	1	−0.173	−0.253	0.06	0.052	−0.236	−0.252	0.043	−0.023	−0.213	−0.084	0.003	−0.088	0.1	0.117	0.075	−0.083	−0.089
Height loss	0.426	−0.173	1	0.28	0.036	0.083	0.275	0.29	−0.002	0.19	0.032	−0.067	0.179	−0.063	−0.021	0.047	0.006	0.183	0.195
Body mass index	0.144	−0.253	0.28	1	−0.066	0.05	0.085	0.05	0	0.042	−0.164	−0.156	0.064	−0.035	−0.152	−0.15	0.28	0.024	0.013
Age at menarche	0.137	0.06	0.036	−0.066	1	−0.327	0.16	−0.032	0.035	0.011	0.084	−0.201	−0.073	0.207	−0.01	0.053	−0.133	0.022	0.032
Menses duration	0.129	0.052	0.083	0.05	−0.327	1	−0.099	0.039	0.029	−0.109	−0.026	−0.004	0.111	−0.168	0.101	0.068	0.055	−0.159	−0.147
Number of labors	0.115	−0.236	0.275	0.085	0.16	−0.099	1	0.535	−0.052	0.013	0.015	−0.162	−0.065	0.001	−0.2	−0.206	0.001	0	0.027
Lactation duration	0.135	−0.252	0.29	0.05	−0.032	0.039	0.535	1	−0.028	−0.046	0.062	−0.012	0.167	−0.029	−0.162	−0.077	0	0.016	0.071
Hip fracture in parents	−0.062	0.043	−0.002	0	0.035	0.029	−0.052	−0.028	1	−0.005	−0.079	0.134	0.122	0.038	−0.048	−0.03	−0.007	0.122	0.128
Number of prior fractures	0.034	−0.023	0.19	0.042	0.011	−0.109	0.013	−0.046	−0.005	1	−0.025	0.017	0.063	0.11	0.113	0.089	−0.07	0.185	0.193
Rheumatoid arthritis	0.072	−.213	0.032	−0.164	0.084	−0.026	0.015	0.062	−0.079	−0.025	1	0.327	0.107	0.019	0.125	0.149	−0.042	−0.006	−0.02
Steroid use	.064	−.084	−0.067	−0.156	−0.201	−0.004	−0.162	−0.012	0.134	0.017	0.327	1	0.093	−0.061	0.101	0.126	0.069	0.085	0.069
Asthma	.024	.003	0.179	0.064	−0.073	0.111	−0.065	0.167	0.122	0.063	0.107	0.093	1	−0.015	0.138	0.225	−0.041	0.156	0.15
Depression	.033	−.088	−0.063	−0.035	0.207	−0.168	0.001	−0.029	0.038	0.11	0.019	−0.061	−0.015	1	0.016	0.02	0.098	0.193	0.169
Calcium supplementation	−0.018	0.1	−0.021	−0.152	−0.01	0.101	−0.2	−0.162	−0.048	0.113	0.125	0.101	0.138	0.016	1	0.889	−0.133	0.036	0.036
Vit D supplementation	0.076	0.117	0.047	−0.15	0.053	0.068	−0.206	−0.077	−0.03	0.089	0.149	0.126	0.225	0.02	0.889	1	−0.153	0.136	0.134
Femoral neck T-score	−0.155	0.075	0.006	0.28	−0.133	0.055	0.001	0	−0.007	−0.07	−0.042	0.069	−0.041	0.098	−0.133	−0.153	1	0.015	0.032
Number of prior falls	0.179	−0.083	0.183	0.024	0.022	−0.159	0	0.016	0.122	0.185	−0.006	0.085	0.156	0.193	0.036	0.136	0.015	1	0.961
Prior falls	0.185	−0.089	0.195	0.013	0.032	−0.147	0.027	0.071	0.128	0.193	−0.02	0.069	0.15	0.169	0.036	0.134	0.032	0.961	1

Correlation   1   −1

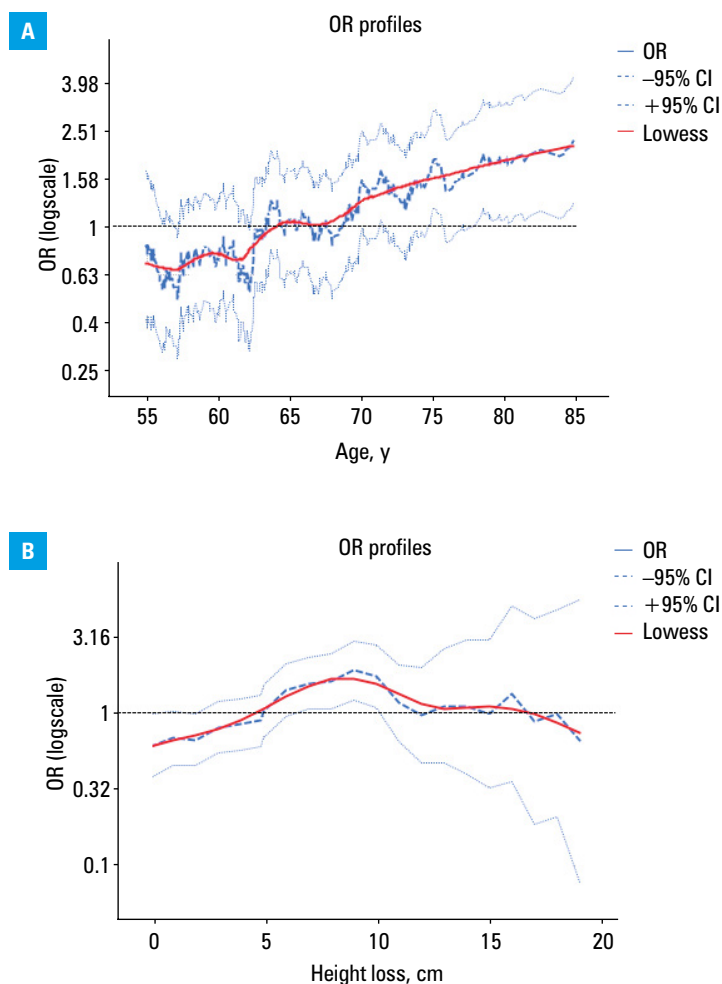**FIGURE 1** Correlation matrix (a heatmap)

FIGURE 2 Odds ratio (OR) profiles for (**A**) age and (**B**) height loss; the red line represents locally weighted scatterplot smoothing.

The shape of the relationship between the continuous features and the predicted variable (fracture occurrence) was analyzed. We checked whether the shape of that relationship was close to a linear one and if it was sufficient to determine the unit odds ratio (OR), or whether it would be more advantageous to divide the predictor variable into categories, that is, to discretize it.

For that purpose, an analysis of the OR profiles was carried out. The graph showing the OR profile for the "Age" variable is shown in **FIGURE 2A**. In this case, an increasing trend can be seen, suggesting that a unit OR can be used.

It was different, for example, in the case of height loss, calculated as the difference between the previous maximum height of a woman and her current height (**FIGURE 2B**). It seemed reasonable to divide this variable into categories. After analyzing the unit changes of OR and its profile, as well as the distribution of height loss, a threshold of 7 cm between the 2 categories of height loss seemed optimal.

Height loss was indicated as a significant predictor of osteoporosis in the univariable logistic analysis but it did not show any significant effect in multivariable logistic regression. Therefore, it was not included in the final model.

The stepwise regression method led to the development of a set of models, each containing a specific set of predictive variables. As suggested by Chakrabarti and Ghosh,[14] in order to select the most appropriate model, the Bayesian information criterion (BIC) should be taken into account, as a tool that can quantify the performance of models and assess their complexity. However, when only 1 model with the lowest BIC is chosen, other models that are equally good or could provide equally useful information, may simply be ignored, therefore, the Bayesian model averaging (BMA) was applied for the set of models in our study.[15,16] It enabled ranking and quantitative comparisons of the competing models and accounted for the uncertainty in variable selection.[17-19] This way, the final set of predictors with high posterior probabilities was determined, and thus the most appropriate model with maximum discriminatory power was selected.

Following the analysis of all the potential predictors, it was decided to include the following variables into the final model: FN T-score, occurrence of fall(s) within the previous 12 months, number of fractures recorded after the age of 40 years, and age.

Age can be treated as a confounding variable (confounder), that is, an external factor implemented into the tested cause-and-effect relationship, which changes its observed covariance by influencing both the explanatory and dependent variables.[20] The inclusion of confounders into the model allows for OR adjustment for the other covariates. Due to the fact that there was a clinically meaningful relationship between age and the analyzed risk factors, as well as the outcome variable, we decided to consider it as a confounding factor.

RESULTS   To predict the 10-year fracture risk, a logistic regression model was developed based on the 4 abovementioned variables. Bootstrap has been recommended as a good resampling method for logistic regression.[21] In the present study, the dataset was split into a 70:30 training:test set ratio. To check whether our model, with the proposed explanatory variables included, was better than the intercept baseline model (ie, a model which does not include any explanatory variables and which is based on the majority category in a dataset), the omnibus tests of model coefficients were used. The presented logistic regression model was significantly better, as evidenced by the χ^2 value of 31.553 (P <0.001).

To study the goodness of fit of the model, the Hosmer–Lemeshow test was performed, showing no significance ($\chi^2 = 4.482$; $P = 0.81$). This test suggested that our logistic regression model was fit to the data and confirmed the relationship between the predictors and the dependent variable, that is, the occurrence of fracture in our study. Details of the examined logistic regression model are shown in **TABLE 3**.

The odds of fracture occurrence depended on the following variables: occurrence of falls within

TABLE 3 Logistic regression model for fracture risk assessment

Predictors	β	SE	95% CI	Wald statistic	OR	*P* value	Bootstrap[a] Bias	Bootstrap[a] BCa 95% CI
Intercept	−3.336	0.952	−5.203 to −1.470	12.274	0.036	<0.001	−0.28	−5.352 to −1.455
Age	0.019	0.015	−0.011 to 0.048	1.572	1.019	0.21	0.000	−0.012 to 0.051
Number of prior fractures	0.437	0.141	0.160–0.714	9.581	1.548	0.002	0.005	0.129–0.747
FN T-score	−0.258	0.123	−0.499 to −0.018	4.421	0.772	0.04	−0.005	−0.532 to −0.011
Prior falls	0.508	0.211	0.094–0.922	5.784	1.662	0.02	0.006	0.080–0.955

a Bootstrap used a bias-corrected and accelerated (BCa) method with 1000 bootstrap samples.

Abbreviations: FN, femoral neck; others, see FIGURE 2
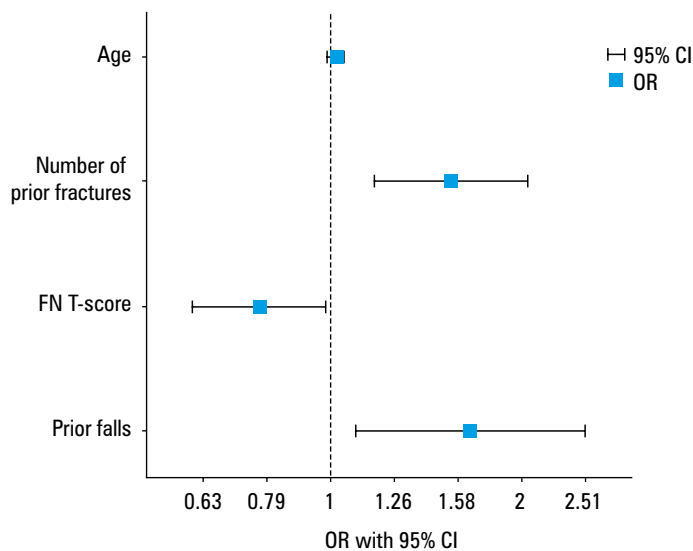
FIGURE 3 Results of the multivariable logistic regression analysis aimed at identifying the fracture risk
Abbreviations: see TABLE 3

the preceding 12 months (referred to as "prior falls" in Tables and Figures; OR, 1.662; 95% CI, 1.099–2.514), number of fractures at the age of 40 years or above (referred to as "number of prior fractures" in Tables and Figures; OR, 1.548; 95% CI, 1.174–2.042), and the FN T-score (OR, 0.772; 95% CI, 0.607–0.983). The ORs were adjusted for the confounding variable of age. The impact of individual predictors on the fracture risk is shown in FIGURE 3.

Taking into consideration the output of the logistic regression analysis and the β coefficients (TABLE 3), the risk of an osteoporotic fracture can be estimated according to the formula shown in FIGURE 4. The algorithm is also available at the www.fracture-risk.pl website.

The prediction accuracy of the proposed model achieved in the test set, expressed by the area under the receiver operating characteristic curve (AUC), is 0.66 (95% CI, 0.604–0.71).

DISCUSSION The development of a 10-year fracture risk algorithm is the major outcome of our study. The applied data components (the population-based epidemiological sample, its size, the long list of features with a potential influence on bone metabolism and status, as well as the long follow-up) allowed us to establish the final algorithm based on reliable data.

Easy access to the algorithm via the Internet and the short time needed to evaluate the risk of fracture make it a user-friendly and straightforward calculator. In clinical practice, simple and reliable tools that facilitate daily work of a physician are needed, rather than complex solutions. In our opinion, the POL-RISK tool fulfils these criteria.

The issue of fracture risk prediction has been the topic of debates of international scientific societies for decades. The possibility to predict the risk of fracture in a patient without fracture history is crucial, as the first fracture significantly increases the risk of subsequent ones. Furthermore, once the fracture risk is established, it may significantly affect the therapeutic management of osteoporosis.

The usual intervention threshold for the 10-year risk of major and any (non-hip) fragility fracture is 10% to 20%, and for hip fractures it is 3%.[22,23] Of note, these recommendations also take economic aspects into account; therefore, they may differ when considered from a purely medical perspective, especially with regard to individual patients. For example, in Australia, in order for a patient to be eligible for medication reimbursement, the risk of any fracture must exceed 26%, while reimbursement is not possible if the risk is below 14%.[3] When the risk is estimated at 14% to 26%, the reimbursement decision depends on the presence of concomitant diseases. The respective values for the hip fracture risk rendering the patient suitable for reimbursement are 3% and 9%[2] (data available at www.fractureriskcalculator.com). In a study performed in a group of 802 men, we showed that the threshold of 7.6% for major fractures and 3.8% for hip fractures, assessed using FRAX, was the best cutoff value, with optimal sensitivity and specificity.[24] Therefore, these thresholds seem to be more appropriate as an indication for treatment initiation than the commonly used values of 10% to 20% and 3% for major and hip fractures, respectively. The issue of which thresholds should be recommended for our algorithm will be the matter of future studies.

$$\text{Risk of fracture} = \frac{1}{1 + e^{-(-3.336 + 0.019 \times \text{Age} + 0.437 \times \text{Number of prior fractures} - 0.258 \times \text{FN T-score} + 0.508 \times \text{Prior falls})}}$$

FIGURE 4 The algorithm for the prediction of the 10-year fracture risk
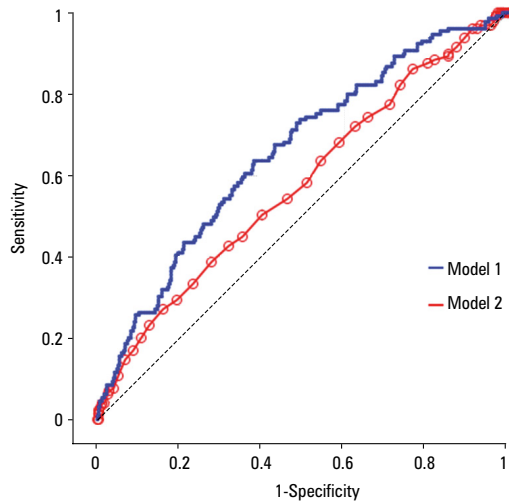Abbreviations: see **TABLE 3**

FIGURE 5 Comparison of receiver operating characteristic curves derived from 2 models; Model 1: age, number of prior fractures, femoral neck T-score, prior falls; Model 2: femoral neck T-score

The design of the present study merits discussion. Some baseline concepts were derived from the analysis of huge databases, as in the case of the studies on FRAX[1] and QFracture.[4] However, we assumed that the most reliable data could be obtained from long-term, prospective observations of large, population-based cohorts. Such an approach was used in the studies carried out by researchers from the Garvan Institute of Medical Research.[2,3] An analysis seeking for significant fracture risk factors is, in our opinion, an optimal way to develop a valuable prediction tool. However, it should be remembered that, while having the proposed recommendations at hand, each patient should be evaluated individually, and the final decision of a physician must not be guided by pure calculations and model solutions.

It is interesting to compare the existing algorithms with respect to fracture risk factors. Generally, they include age, previous fractures, and BMD values, usually expressed by T-score.[1,3-5] In the Garvan and POL-RISK models, the number of fractures is also taken into account, while it is not included in the FRAX method.

The history of falls as a fracture risk factor is included both in the Garvan algorithm and the POL-RISK calculator, but the Australian method also takes into account the number of falls during the previous 12 months. The FRAX tool does not include any information about falls, however, several other risk factors modify the final fracture probability level. The broadest spectrum of possible risk factors is included in the QFracture tool.

Recently, Beaudoin et al[25] published a review and meta-analysis focused on the performance of predictive tools used to identify individuals at risk of nontraumatic fracture. The authors identified 53 validation studies assessing the discriminative ability of 14 tools. Due to the small number of studies on some prediction tools, only the FRAX, Garvan, and QFracture methods were compared using meta-regression models. The authors concluded that QFracture, FRAX with BMD, and Garvan with BMD had the best discriminative performance for predicting fracture. We compared the AUC values reported in these models with the value obtained in our study. We found that QFracture[4] had the best discriminative ability to predict major fractures (AUC = 0.77), and the FRAX and Garvan methods both presented an AUC of 0.72. These AUC values are higher than the one obtained in our study (AUC = 0.66); however, it should be emphasized that they were established in different study cohorts. Therefore, a direct comparison between them does not seem feasible or valid. The application of the Garvan and FRAX models to the dataset used in our follow-up study led to worse outcomes than those mentioned above (AUC = 0.64 and AUC = 0.65, respectively, data not shown). It is also worth mentioning that although the AUC in our study seems not very high, the AUC calculated solely on the basis of the FN T-score, which, together with age, is considered the most important factor for the diagnosis of osteoporosis, was only 0.58 (95% CI, 0.523–0.634) (**FIGURE 5**). Therefore, it can be concluded that the proposed model is more accurate than the currently used fracture risk models and may help identify patients who are in need of osteoporosis therapy or diagnostics. More detailed comparisons between our model and other popular fracture risk calculators will be the subject of a separate publication. Of note, in the current study, a Lunar densitometer was used, and the accuracy of fracture risk assessment using BMD values obtained with other devices warrants further investigation.

The idea of developing a fracture risk calculator is still very popular. On one hand, continuous validation of the available tools and attempts to identify risk factors potentially specific to regions or ethnic groups are being postulated. On the other hand, in light of the development of bone tissue assessment methods other than densitometry, such as radiofrequency echographic multispectrometry,[26] we can also expect algorithms based on diagnostic tools alternative to dual energy X-ray absorptiometry.

Certain limitations of our study need to be mentioned. First, the follow-up only targeted women. Second, spine radiographs were not available for all participants, so some clinically silent

vertebral fractures might have been missed. Also, the number of hip fractures was too small to allow for the development of a separate algorithm for the prediction of hip fracture risk. However, the large population-based sample, the number of factors with potential influence on fracture risk, and the long-term prospective follow-up made it possible to develop a new fracture risk prediction tool. We plan to analyze and verify the predictive capability of the presented fracture risk algorithm in other populations.

Prospective follow-up seems to be the most reliable approach for the identification of factors determining long-term outcomes; however, due to the fact that a simple and reliable risk assessment tool is urgently needed in daily clinical practice, we plan to carry out a retrospective, longitudinal study in the near future, involving the group of patients previously analyzed in the GO-Study.[27] Such a study would be helpful to verify the accuracy of fracture prediction of the POL-RISK tool.

Due to the long follow-up and the difference between the size of the group at the time of recruitment and at the end of the study, the dropout phenomenon should be commented on. The total dropout rate over the 10 years was 34.5%, which is satisfactory taking into account the length of the follow-up and the age of the participants at the baseline. Nevertheless, it should be taken into account that the drop-out process may not be fully random and may lead to selective elimination of individuals with certain clinical features. Therefore, we decided to conduct an atypical analysis of the basic clinical characteristics of women who did not complete the study. Higher age in the subgroup lost to follow-up was clearly expected due to higher mortality among older study participants and increasing difficulty in maintaining contact with aging people. Apart from this parameter, the subgroups did not differ in any of the most important factors affecting the fracture risk. Therefore, it can be assumed that the group participating in the final analysis retained its character, representative of the population, which was warranted in the group recruited for the study.

To conclude, we developed a 10-year fracture risk prediction algorithm. Its use and outcomes should be followed by appropriate therapeutic regimens to reduce the number of future fragility fractures, thus decreasing the financial burden on health care systems.

SUPPLEMENTARY MATERIAL

Supplementary material is available at www.mp.pl/paim.

ARTICLE INFORMATION

REFERENCES

1 Kanis JA, Johnell O, Oden A, et al. FRAX™ and the assessment of fracture probability in men and women from the UK. Osteoporos Int. 2008; 19: 385-397.

2 Nguyen ND, Frost SA, Centre JR, et al. Development of a nomogram for individualizing hip fracture risk in men and women. Osteoporos Int. 2007; 18: 1109-1117.

3 Nguyen ND, Frost SA, Centre JR, et al. Development of prognostic nomograms for individualizing 5-year and 10-year fracture risks. Osteoporos Int. 2008; 19: 1431-1444.

4 Hippisley-Cox J, Coupland C. Predicting risk of osteoporotic fracture in men and women in England and Wales: prospective derivation ad validation of QfractureScores. BMJ. 2009; 339: b4229.

5 Pluskiewicz W, Adamczyk P, Czekajło A, et al. Epidemiological data on osteoporosis in women from the RAC-OST-POL study. J Clin Densitom. 2012; 15: 308-314.

6 Adamczyk P, Werner A, Bach M, et al. Risk factors for fractures identified in the algorithm developed in 5-year follow-up of postmenopausal women from RAC-OST-POL study. J Clin Densitom. 2018; 21: 213-219.

7 Zagórski P, Tabor E, Martela-Tomaszek K, et al. Five-year fracture risk assessment in postmenopausal women, using both the POL-RISK calculator and the Garvan nomogram: the Silesia Osteo Active Study. Arch Osteoporos. 2021; 16: 32.

8 Kostrzewa D, Brzeski R. The data dimensionality reduction and features weighting in the classification process using Forest Optimization Algorithm. In: Huk M, Maleszka M, Szczerbicki E (eds). Intelligent Information and Database Systems: Recent Developments. ACIIDS 2019. Studies in Computational Intelligence. Cham: Springer; 2020: 830.

9 Bach M, Werner A. Cost-sensitive feature selection for class imbalance problem. In: Proceedings of 38th International Conference on Information Systems Architecture and Technology - ISAT 2017. ISAT 2017. Advances in Intelligent Systems and Computing. Cham: Springer; 2018: 655.

10 Vatcheva KP, Lee M, McCormick JB, Rahbar MH. Multicollinearity in regression analyses conducted in epidemiologic studies. Epidemiology (Sunnyvale). 2016; 6: 227.

11 Kutner M, Nachtsheim C, Neter J. Applied Linear Statistical Models, 4th edition. New York, NY: McGraw-Hill/Irwin; 2004.

12 Berry WD, Feldman S. Multiple Regression in Practice (Quantitative Applications in the Social Sciences) Thousand Oaks, CA: SAGE Publications; 1985.

13 Belsley DA. Conditioning Diagnostics: Collinearity and Weak Data in Regression. 1st. John Wiley & Sons; 1991.

14 Chakrabarti A, Ghosh JK. AIC, BIC and recent advances in model selection. Philosophy of Statistics. 2011; 7: 583-605.

15 Yang F, Qingfeng L, Qingsong Y, Guoqing Z. Model averaging for nonlinear regression models. Journal of Business & Economic Statistics. 2022; 40: 785-798.

16 Murphy TE, Tsang SW, Leo-Summers LS, et al. Bayesian model averaging for selection of a risk prediction model for death within thirty days of discharge: the SILVER-AMI study. Int J Stat Med Res. 2019; 8: 1-7.

17 Piironen J, Vehtari A. Comparison of Bayesian predictive methods for model selection. Stat Comput. 2017; 27: 711-735.

18 Dormann CF, Calabrese JM, Guillera-Arroita G, et al. Model averaging in ecology: a review of Bayesian, information-theoretic, and tactical approaches for predictive inference. Ecol Monogr. 2018; 88: 485-504.

19 Galipaud, M, Gillingham MAF, Dechaume-Moncharmont, FX. A farewell to the sum of Akaike weights: the benefits of alternative metrics for variable importance estimations in model selection. Methods Ecol Evol. 2017; 8: 1668-1678.

20   Hernán MA, Hernández-Díaz S, Werler MM, Mitchell AA. Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology. Am J Epidemiol. 2002; 155: 176-184.

21   Vrigazova B, Ivanov I. The bootstrap procedure in classification problem. International Journal of Data Mining, Modelling and Management. 2020; 12: 428-446.

22   Lorenc R, Głuszko P, Karczmarewicz E, et al. Recommendations for diagnostic and therapeutic management in osteoporosis – update 2013 [in Polish]. Medycyna Praktyczna – Reumatologia. 2013; special edition 1/2013: 3-44.

23   Camacho PM, Petak SM, Binkley N, et al. American Association of Clinical Endocrinologists and American College of Endocrinology clinical practice guidelines for the diagnosis and treatment of postmenopausal osteoporosis. Endocrine Practice. 2016; 22: 1-42.

24   Pluskiewicz W, Adamczyk P, Franek E, et al. FRAX calculator and Garvan nomogram in male osteoporotic population. Aging Male. 2014; 17: 174-182.

25   Beaudoin C, Moore L, Gagné M, et al. Performance of predictive tools to identify individuals at risk of non-traumatic fracture: a systematic review, meta-analysis, and meta-regression. Osteoporos Int. 2019; 30: 721-740.

26   Nowakowska-Płaza A, Wroński J, Płaza M, et al. Diagnostic agreement between radiofrequency echographic multispectrometry and dual-energy X-ray absorptiometry in the assessment of osteoporosis in a Polish group of patients. Pol Arch Intern Med. 2021; 131: 840-847.

27   Pluskiewicz W, Adamczyk P, Drozdzowska B. Height loss in postmenopausal women – do we need more for fracture risk assessment? Results from GO Study. Osteoporos Int. 2021; 32: 2043-2049.