Journal of Clinical Densitometry: Assessment & Management of Musculoskeletal Health, vol. 21, no. 2, 213–219, 2018

1094-6950/21:213–219/$36.00
https://doi.org/10.1016/j.jocd.2017.07.005

Risk Factors for Fractures Identified in the Algorithm Developed in 5-Year Follow-Up of Postmenopausal Women From RAC-OST-POL Study

P. Adamczyk,[*,1] *A. Werner,*[2] *M. Bach,*[2] *J. Żywiec,*[3] *A. Czekajło,*[4] *W. Grzeszczak,*[3] *B. Drozdzowska,*[5] *and W. Pluskiewicz*[6]

[1]*Department of Pediatrics, School of Medicine with the Division of Dentistry in Zabrze, Medical University of Silesia in Katowice, Poland;* [2]*Faculty of Automatic Control, Electronics and Computer Science, Silesian University of Technology, Gliwice, Poland;* [3]*Department of Internal Medicine, Diabetology and Nephrology, School of Medicine with the Division of Dentistry in Zabrze, Medical University of Silesia in Katowice, Poland;* [4]*Dialysis Station Fresenius NefroCare, Wodzisław, Poland;* [5]*Department of Pathomorphology, School of Medicine with the Division of Dentistry in Zabrze, Medical University of Silesia in Katowice, Poland; and* [6]*Department of Internal Medicine, Diabetology and Nephrology—Metabolic Bone Diseases Unit, School of Medicine with the Division of Dentistry in Zabrze, Medical University of Silesia in Katowice, Poland*

Abstract

The aim of the study was to establish factors with an impact on fracture risk and to develop an algorithm to predict osteoporotic fracture. A total of 978 postmenopausal women from the epidemiological, population-based RAC-OST-POL study with a mean age of 65.7 ± 7.3 years were enrolled. At baseline, bone mineral density at hip and clinical risk factors for fracture were collected. Afterward, each person was asked annually on fracture incidence in the 5-year follow-up. Finally, data for complete 5-year observation were gathered for the group of 802 patients. During the follow-up, 92 osteoporotic fractures occurred in 78 women. The most common fracture site was the forearm (n = 45). The following baseline factors were found as significant for fracture incidence: femoral neck bone mineral density, prior fractures, steroid use, falls within previous 12 months, and height. Fracture risk was predicted by the following formula: $Risk\ of\ \ fracture\ incidence = \frac{1}{1+e^{-\left(\begin{array}{l}-9.899+1.077*STEROIDS+0.681\\ *PRIORFALLS+0.611*PRIORFRACTURES\\ -0.483*FN\ Tscore+0.042*HEIGHT\end{array}\right)}}$. In our current longitudinal study, an algorithm predicting fracture occurrence over a period of 5 years was developed. It may find application in daily medical practice.

Key Words: Bayesian Model Averaging; follow-up; fracture risk; logistic regression; postmenopausal women.

Introduction

Osteoporosis is one of the most important health problems in the elderly. Patients with osteoporosis do not usually have any clinical symptoms. That is why osteoporosis is also called a "silent epidemic." The most important consequences of osteoporosis are fractures that can be caused even by minimal trauma, for example, a fall from standing height. One may expect that such fracture will occur in approximately 40% of postmenopausal women *(1)*.

Received 04/20/17; Revised 07/7/17; Accepted 07/21/17.
P. Adamczyk, A. Werner, M. Bach, J. Żywiec, P. A. Czekajło, W. Grzeszczak, B. Drozdzowska, and W. Pluskiewicz declare that they have no conflict of interest related to this article.
*Address correspondence to: Piotr Adamczyk, MD, Department of Pediatrics in Zabrze, 3 Maja 13/15 street, 41-800 Zabrze, Poland. E-mail: adamczyk@konto.pl

According to the World Health Organization, there are a number of osteoporosis risk factors, namely, prior fractures, the level of physical activity, smoking, alcohol intake, a family history of fracture, age, and many others. Some new methods of assessing fracture risk on the basis of various factors, including those abovementioned, in the last decade, have been developed. Among them are FRAX *(2)*, the method proposed by Garvan Institute *(3,4)*, and QFracture *(5)*. These methods are dedicated to estimate an individual's risk of osteoporotic fracture occurrence over the next 5 (Garvan) or 10 years (FRAX, Garvan, QFracture). The risk given by FRAX is modified by expected life duration, and in fact it expresses fracture probability. The established fracture probability according to FRAX results is proposed as a threshold for the beginning of pharmacologic treatment *(6,7)*. However, it should be underlined that because of several reasons, a fracture risk can be different in various populations, and the model of fracture prediction derived in 1 country not necessarily expresses a risk in another country properly. Therefore, studies showing fracture risk for each society are extremely necessary.

The RAC-OST-POL is such a kind of study that was performed on Polish postmenopausal women in the year 2010. The matters of interest were, among others, nutrition *(8,9)*, functional status *(10)*, the role of education, marital status, kind of job or place of living *(11)*, and vision impairment *(12)*. In this epidemiological, population-based study, the osteoporotic fractures were observed in 28% of all women *(13)* older than 40 years. Besides, in our recent paper, fracture incidences in 4-year follow-up were analyzed in regard to fracture risk (Garvan) and probability (FRAX) *(14)*.

In the current research, we developed a model of fracture risk assessment, and the following detailed aims of the study were defined:

(1) determination of risk factors on the basis of the RAC-OST-POL study with the use of various techniques and statistical models,
(2) creation of the new model for predicting fracture risk,
(3) evaluation of the proposed model for predicting fracture risk by estimating its prediction accuracy.

Materials and Methods

The RAC-OST-POL study is an epidemiological, population-based program designed to reveal data related to postmenopausal osteoporosis. Women who make up the cohort of the RAC-OST-POL study were randomly selected from local population of the whole Racibórz district in Southern Poland *(2)*.

During the 5-year follow-up, phone calls were performed in May 2011–2015, and all fractures of nontraumatic origin were noted. Each patient was asked for confirmation of the fracture by her doctor, and only confirmed events were included in the study. The initial cohort consisted of 978 women. In the case of 176 women, data were censored because of loss of contact, as 131 women did not respond to either phone call or our letters probably because of phone number or postal address change, 40 women died, and 5 refused to cooperate. Finally, data for all 5 observation points were available on the group of 802 patients.

Table 1
Baseline Characteristics of RAC-OST-POL Subjects Enrolled to the 5-Year Longitudinal Observation, n = 802

Parameter	Mean	SD	Median
Age (y)	65.52	7.23	64.49
Height (cm)	156.32	5.73	157.00
Weight (kg)	74.58	13.80	73.00
BMI (kg/m^2)	30.53	5.44	29.71
Menarche (y)	13.97	1.67	14.00
Menopause (y)	49.25	4.88	50.00
FN T-score	–1.24	0.91	–1.30

For all study participants, various data suspected as potential osteoporosis or fracture risk factors were collected *(2)*.

Bone mineral density (BMD) for both femoral neck (FN BMD) and total hip (TH BMD) were established using Lunar DPX (GE, Madison, USA).

Baseline data of the RAC-OST-POL study sub-cohort enrolled to the longitudinal observation is presented in Table 1.

At the beginning of the RAC-OST-POL study, the information about previous 373 fractures that happened over the age of 40 years in 286 patients was collected. Meanwhile, 92 new fractures occurred (hip, 4; spine, 15; forearm, 45; arm, 6; lower leg, 10; rib, 3; foot, 5; clavicle, 4) in n = 78 women during the 5-year follow-up. We present data of all fractures noted over the period of observation, but in the statistical analyses described below, fracture prediction concerns only the first fracture incidence in follow-up.

Statistics

For all calculations, the $p < 0.05$ was taken as a cutoff point of statistical significance. The analyses and computations were performed in 3 main steps using the particular software environments and tools listed in the corresponding paragraphs.

Step 1. Analysis of Risk Factors for Osteoporosis

In our original database, each patient was initially described by nearly 200 variables, but the preliminary reduction in dimensions of source database was done for the purpose of data quality improvement. Some of the features were eliminated because they were redundant or insignificant and did not affect the existence or absence of the fracture (e.g., name, marital status, and so forth). It was accomplished with Waikato Environment for Knowledge Analysis software Weka 3.7.13, developed at the University of Waikato, New Zealand *(15)*, and software environment for statistical computing and graphics R project *(16)*.

More information about this stage of data preparation and the detailed results of the chosen methods are included in our previous study *(17)*.

The importance of each variable was evaluated by the use of the following techniques:

- individual predictive ability on the basis of the degree of redundancy between all features,
- gain ratio (GainRatio) criterion,
- information gain (InfoGain),
- the Gini index.

The subset of features found in the step of feature selection was the source for Bayesian Model Averaging. This analysis was performed to ascertain the most appropriate model with maximum discriminatory power.

Step 2. Logistic Regression Analysis

The statistical analysis was carried out with the professional statistical environment Statistica version 12 by StatSoft Inc (Tulsa, OK). Besides, the data were analyzed with the statistical computing software copyrighted by PQStat Software, Poznań, Poland, version 1.6.2 *(18)*.

The models of logistic regression were built and it allowed to determine relevant variables and reveal their impact on the presence or absence of fractures. Each model was generated with the usage of the bootstrap method, which can reduce effects of data distribution bias. This method relies on random sampling with replacement. It computes the statistic from the resampled data, and aggregates over multiple realizations (resamples) of the data.

The verification of the model usefulness was conducted by performing the Omnibus Tests of Model Coefficients based on the chi-square test. The test allows to check whether at least one of the predictors from the given set of variables is the significant factor of the estimated probability. The value of probability (p value) less than 0.05 indicates that the examined model outperforms the random model, that is, the model of accidental anticipation of a fracture. To evaluate the logistic regression model with regard to goodness-of-fit statistics, the Hosmer-Lemeshow test was used.

To gain the possibility of more thorough analyses of the particular categories and to clarify the impact of a given variable on a chance of fracture incidence, there were built models in which chosen variables were split into dummy ones. In the analyses, with the usage of dummy variables, the odds ratio was interpreted as the odds of the fracture occurrence calculated with regard to the specified reference category.

FN T-score was divided into 3 categories with different FN BMD, namely, *no risk* (FN T-score ≥ –1), *osteopenia* (–1 > FN T-score ≥ –2.5), and *osteoporosis* (FN T-score < –2.5). *No risk* constituted the reference category. Moreover, the variable PRIOR FRACTURES[1] was replaced by NUM FRACTURES, which was categorized as follows: without fractures (group 0), 1 or 2 fractures (group 1), and more than 2 fractures (group 3). The reference category was group 0.

[1]PRIOR FRACTURES variable indicated only the presence or absence of prior fracture, whereas NUM FRACTURES contained information on specific number of earlier fractures.

Step 3. Assessing of the Proposed Model

To assess the prediction accuracy of the proposed model, the receiver operating characteristic (ROC) was studied, as well as area under the curve (AUC) calculation with the usage of DeLong method. ROC curve was applied to compare the diagnostic strength of the proposed model and FN T-score, which is the most reliable gold standard of osteoporosis. Besides, the optimal cutoff point with the best accuracy to predict fractures was determined using both ROC curve analysis and the Youden index.

Results

To gain transparency, all results are presented in 3 points with regard to the specified stages of statistical analysis.

Step 1. Analysis of Risk Factors for Osteoporosis

As a result of different preprocessing strategies, the initial dimension of the database was reduced. Finally, 25 variables suggested by majority of filtering methods were chosen for further analyses: age, height and body mass index, results of Timed Up and Go mobility test, age of first menarche, the duration of menses (in years) and lactation (in months), the number of labors, the presence of comorbidities such as rheumatoid arthritis and diabetes type 1, steroids administration, smoking, parental hip fracture, BMD for femoral neck determined by dual-energy X-ray absorptiometry and reported in 2 numbers: T-score and Z-score, incidents of prior fractures and their number, incidents of prior falls within the last 12 months, and the indicator whether the patient meets the "gold standard" criteria for osteoporosis.

The Bayesian Model Average analysis conducted on the basis of our data indicated the several optimal models for fracture risk prediction. The final model considered only 5 variables: height, steroids administration, T-score of BMD of femoral neck (FN T-score), incidents of prior fractures, and incidents of prior falls, reported during the first medical examination (Table 2).

Table 2
Osteoporosis Risk Factors Included in the Final Model of BMA

Risk factor	Data type	
Steroids administration	0 = No	1 = Yes
Height	Quantitative (cm)	
Prior fractures	0 = No	1 = Yes
Prior falls	0 = No	1 = Yes
FN T-score	Quantitative	

Abbr: BMA, Bayesian Model Averaging.

Table 3
The Model of Logistic Regression for Risk of Fractures

							Bootstrap[a]			
									BCa 95% confidence interval	
	β[b]	S.E.	−95% CI	+95% CI	Wald stat.	Odds ratio (OR)	Bias	*p* value	Lower	Upper
Constant	−9.899	3.518	−16.794	−3.004	7.919	0.000	−0.337	0.001	−15.765	−4.572
Steroids	1.077	0.413	0.266	1.887	6.780	2.935	−0.024	0.009	0.091	1.136
Prior falls	0.681	0.250	0.190	1.172	7.390	1.975	0.011	0.010	0.181	1.230
Prior fractures	0.611	0.253	0.116	1.107	5.844	1.843	0.008	0.015	−0.815	−0.193
FN T-score	−0.483	0.152	−0.782	−0.184	10.047	0.617	−0.012	0.006	0.076	1.883
Height	0.042	0.022	−0.001	0.085	3.611	1.043	0.002	0.030	0.001	0.086

[a]Bootstrap was performed with bias-corrected and accelerated (BCa) method. Its results were based on 5000 bootstrap samples.
[b]Risk of fracture incidence for a patient is calculated by the use of the following formula: 1/(1 + e^(−(−β + 1.077*STEROIDS + 0.681*PRIORFALLS + 0.611*PRIORFRACTURES − 0.483*FN T score + 0.042*HEIGHT))).

Step 2. The Model of Regression Analysis

To predict 5-year fracture risk, the abovementioned features were considered. The logistic regression model was constructed and evaluated to check which of these variables had a significant influence on the occurrence of the fracture.

The results of the Omnibus Tests of Model Coefficients based on the chi-square test showed that the created logistic regression model was significant (chi-square = 32.895, *p* value = 0.0).

To study the goodness of fit of the model, the Hosmer-Lemeshow test was performed and showed no significance (chi-square = 7.969; *p* = 0.436). This test revealed that our logistic regression model fitted to the data and confirmed relationship between the predictors and the dependent variable, which in our study was the fact of fracture incidence. The details of the examined model of logistic regression are shown in Table 3.

The odds of fracture occurrence depends on the variables in the manner described by the odds ratio (OR):

- *Steroids* variable: OR [95% confidence interval {CI}] = 2.935 [1.305; 6.599]—the odds of osteoporotic fracture increases nearly 3 times for individuals who took steroids.
- *Prior falls* variable: OR [95%CI] = 1.975 [1.209; 3.227]—the odds of osteoporotic fracture for the individuals who experienced earlier falls is approximately twice higher than for the ones with no falls.
- *Prior fractures* variable: OR [95%CI] = 1.843 [1.123; 3.024]—the odds of the occurrence of the fractures for the individuals who underwent fractures is 1.84 times greater than for the ones without prior fractures.
- *Height* variable: OR [95%CI] = 1.041 [1.001; 1.088]—when the height of individuals increases by 1 cm, the odds of new fractures increases 1.04 times.
- FN T-score variable: OR [95%CI] = 0.617 [0.458; 0.832]—the higher the FN BMD T-score value, the smaller odds of the occurrence of the osteoporotic fracture.

After discretizing the FN T-score variable, the odds ratio for *osteopenia* category is OR [95%CI] = 1.54 [0.87; 2.72], whereas for *osteoporosis* is OR [95%CI] = 5.57 [2.56; 12.14]. That means that the individuals who belong to the category *osteopenia* have more than 1.5 times greater chance of subsequent fracture than those belonging to category *no risk*. For people who belong to the category *osteoporosis* with FN T-score <−2.5, the odds of fracture steeply grows up to 5.57 times (Fig. 1).

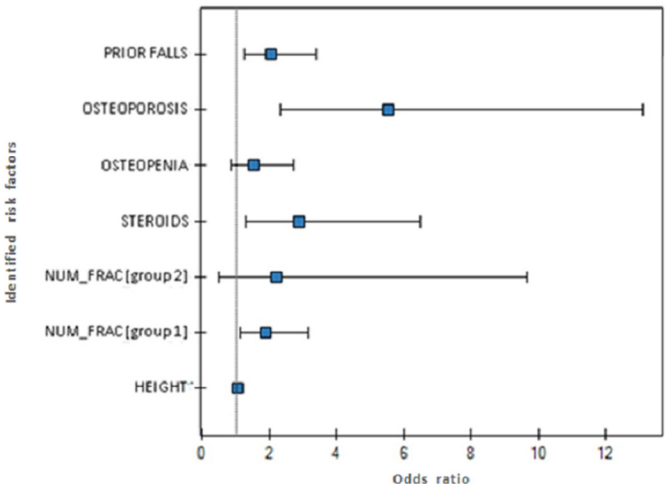

Fig. 1. Impact of discretized variables on the occurrence of fracture. X-axis: the multiplicity estimates of confidence intervals (CI) for odds ratios (OR). Legend: I 95%CI, ■, OR.

After discretizing the number of fractures (NUM FRACTURES), OR [95%CI] = 1.8 [1.09; 2.99] for group 1 (with 1–2 fractures) and OR [95%CI] = 2.48 [0.59; 10.33] for group 2 (more than 2 fractures). Thus, it can be concluded that the individuals with 1 or 2 fractures have more than 1.8 times greater chance of subsequent fracture than the individuals without any fracture. Moreover, for those with more than 2 fractures, the chance of subsequent fracture rises up to more than 2.4 times than those from group 0 (Fig. 1).

The use of logistic regression allows to model the probabilities of a response variable (Y) as a function of 1 or more predictors (X_i). It can be estimated by the formula:

$$P(Y=1)=\frac{1}{1+e^{-\left(\beta_0+\sum_{i=1}^{k}\beta_i X_i\right)}}. \quad (1)$$

In the presented study, X_i relates to proposed 5 risk factors for fractures. The life expectancy is not taken into consideration, so by substituting the β_i coefficients from formula (1), with the values taken from the output of the logistic regression analysis from Table 3, the risk of fracture incidence can be estimated for the particular individuals as follows:

$$\textit{Risk of fracture incidence} = \frac{1}{1+e^{-\left(\begin{array}{l}-9.899+1.077*STEROIDS+0.681\\ *PRIORFALLS+0.611*PRIORFRACTURES\\ -0.483*FN\ T\ score+0.042*HEIGHT\end{array}\right)}}. \quad (2)$$

The value of $e^{\beta 0}$ is interpreted as a chance of the fracture occurrence in a reference group.

As an example, 2 patients from the RAC-OST-POL study cohort can be chosen, and their risks can be calculated on the basis of the specific values of the variables. Taking corresponding values of the variables for patient 1 (steroids = 1, prior falls = 1, prior fractures = 1, FN T-score = –1.2, and height = 158[cm]), the fracture risk is estimated as 0.422. By contrast, the approximate fracture risk for patient 2 with FN T-score = –0.5, height = 163 cm, and without prior fractures, prior falls, or steroids administration equals 0.057. The acquired data can be expressed in percentage by multiplying it by 100. Consequently, it can be stated that the estimated risk of fracture incidence after 5 years for patient 1 is 42.2%, whereas for the second patient it is only 5.7%.

Step 3. Assessment of the Proposed Model

For the proposed model, achieved prediction accuracy expressed by the parameter AUC is 0.69 [95% CI 0.626–0.753]. AUC seems to be not very high; however, AUC value calculated based exclusively on the FN T-score, which is the gold standard for diagnosis of osteoporosis, equaled only 0.58 [95% CI 0.517–0.654] (Fig. 2). Thus, it can be concluded that the proposed model appears to be better than the currently accepted standard and it helps better to identify groups of patients who should be treated or precisely diagnosed in a specialized clinic.

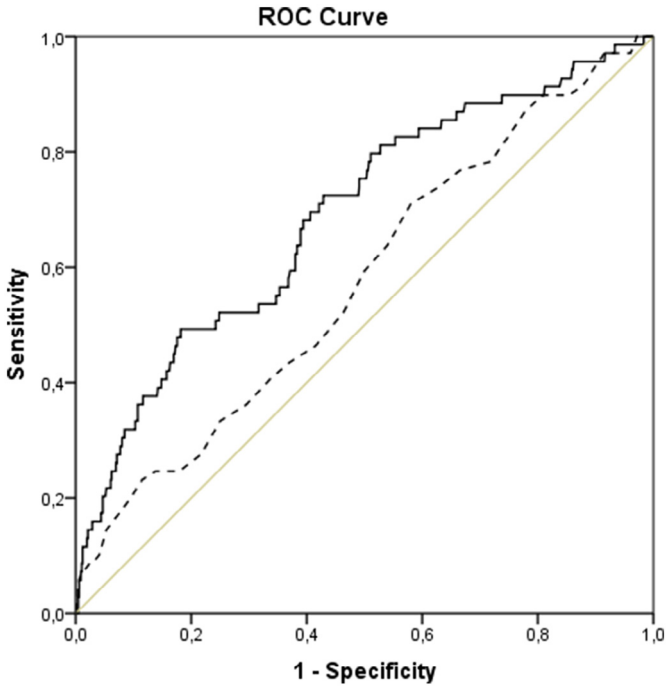

Fig. 2. Comparison of ROC curves. Source of the curve: reference line; - - FN T-score, — REGR-RAC-OST-POL.

According to Cramer's rule of optimal limit for imbalanced data *(19)*, if the share of $Y = 1$ in a sample is δ, then Y is predicted to be 1 for new observations with the empirical risk ≥δ or 0, otherwise. In our case, Y = 1 meant the share of people who underwent a bone fracture after 2010. For the analyzed data, δ was 0.1. Thus, according to Cramer's rule, the cutoff point was 0.1 (sensitivity = 0.615, specificity = 0.621), but, for example, the Youden index determined the slightly lower cutoff value: 0.091 (sensitivity = 0.731, specificity = 0.567).

Considering the abovementioned value of δ, patient 1 analyzed in the previous step can be ultimately regarded as the person with the high risk of fracture incidence (estimated risk of fracture is 0.422, so the value is >0.091). The validity of such a conclusion has been empirically confirmed. In the case of this patient, such a fracture really appeared in 5 years after preliminary examination. In contrast, patient 2 has lower estimated risk of fracture incidence. It equals 0.057 which is below the value of cutoff point (<0.091) and actually it was verified that this patient has not suffered from any fracture until now.

Discussion

The crucial issue in management of osteoporotic patients concerns the proper estimation of fracture risk. Therefore, in this study, the new algorithm assessing the risk in a period of 5 years was developed.

As widely known, fracture risk is dependent on 2 main causes: bone status and functional status leading to falls. The bone status is related to several modifiable and unmodifiable factors, like age, sex, genetics, comorbidities, used medications, nutrition, smoking, and physical activity. The measurement of BMD expresses the current bone status and is considered as one of the most important risk factors for a fracture. In our algorithm, well-known and established risk factors for fracture were included, such as low BMD value, falls within 12 previous months, steroid use, height, and prior fractures. Comparing the factors with those from the other calculators such as FRAX or Garvan, it can be noticed that the prior fractures is a widely accepted factor for subsequent fractures, although it is not a rule (e.g., in QFracture tool, this feature is not taken into account). One may expect that increasing age would play an important role as a risk factor, but in this study, age was verified to be statistically insignificant. It is not easy to explain this surprising observation. We suppose that because of the relatively short duration of follow-up, the age is not as significant as expected. The longer follow-up period up to 10 years might present greater importance of the age factor. In FRAX algorithm, age is included as 1 of the 12 factors, but, for example, falls are not taken into consideration. On the other hand, the Garvan algorithm comprises not only age and prior fractures, but also falls, T-score, and FN BMD. This differences may result from a number of the reasons, but additionally one should also remember that FRAX is not based on longitudinal observation, whereas the Garvan algorithm is based on long-term observation of Australian population in Dubbo. Our data are also derived from longitudinal observation of population-based sample, which confirms the importance of current results. Another approach was used in creating QFracture scores *(5)*. The authors gathered an incredibly great number of patients' records, 2.4 million patients aged 30–85 years. BMD and prior fractures were not included as the potential risk factors. This algorithm comprises several factors including age, history of falls, and steroid use. QFracture scores was validated in prospective observation, and this study revealed a good performance for osteoporotic fracture *(20)*. AUC for women exceeds 0.8, which is a better value than AUC of 0.69 noted in the current study.

The relatively low value of the AUC obtained in our model could be probably attributed to the fact that the patients with bone fractures after 2010 represented only about 10% of the study cohort, which means that initial dataset was imbalanced. In this case, the improvement in the quality of prediction of the examined model can be achieved by removal of skewed class distribution. Our paper *(21)* contains more detailed discussion about the methods that allow to solve this problem. Choosing the appropriate methods and level of balancing allow to obtain the higher value of AUC (i.e., AUC = 0.741 for 25% of the patients with bone fractures within the dataset).

Our study has some strengths, but there are some limitations as well. To our strengths belong the fact that the studied cohort was selected from a female population older than 55 years of the whole district and it can be considered as the representative sample. At baseline, we gathered many important clinical features possibly related to osteoporosis and fractures. The bone status was established using dual-energy X-ray absorptiometry at proximal femur. During the period of observation, the number of fractures was high enough to perform reliable statistical analysis.

The next strength of our algorithm is the fact that we gathered the representative data from the urban district from the south of Poland. It is known that factors that influence the possibility of the osteoporotic illness may be various for different regions of the world because they can depend on the levels of technological advancement, climate, economic conditions, and so on; therefore, an individual approach seemed to be valuable. Moreover, as we already know, no osteoporosis calculator dedicated to the population of Poland has been developed, even though such a tool would be useful in the daily practice of Polish physicians.

As limitations of the study, we can mention the lack of direct confirmation of reported fractures (fractures were recorded based on information given by responders during phone calls after confirmation by general practitioners taking care of studied subjects) and the drop-out phenomenon. The full 5-year longitudinal observation was available in 802 of 978 initially recruited subjects, and only these data were taken into consideration. This means that we noted the drop-out ratio at 18%, which is, however, acceptable for such long follow-up period.

The algorithm derived in the current study requires further validation. For that moment, such validation is not possible to be performed because alternative methods are not available. First, FRAX presents fracture probability and we calculated fracture risk. Second, FRAX estimates major osteoporotic fractures whereas our algorithm expresses the risk of any fracture. Third, FRAX shows fracture probability for a decade and our algorithm is limited to the period of 5 years. Fourth, although Garvan fracture risk calculator also has a 5-year variant, the validation of our finding in comparison with data derived in Australian population cannot be rationally performed in our opinion. We consider that the only way to achieve real validation is further, prospective observation of great populations in regard to concordance between baseline value of fracture prediction presented by our algorithm and fractures observed in longitudinal study.

Concluding, in the current longitudinal study, the algorithm predicting fracture incidence over the period of 5 years was developed. Even though our current algorithm is based on the short 5-year follow-up, it can be used in daily work with patients. In further research, we plan to upgrade the algorithm after 10-year follow-up.

References

1. Melton LJ, Atkinson EJ, O'Fallon WM, et al. 1991 Long term fracture risk prediction with bone mineral measurements made at various skeletal sites. J Bone Miner Res 6(Suppl 1):S136.

2. Kanis JA, Johnell O, Oden A, et al. 2008 FRAX™ and the assessment of fracture probability in men and women from the UK. Osteoporos Int 19:385–397.
3. Nguyen ND, Frost SA, Centre JR, et al. 2007 Development of a nomogram for individualizing hip fracture risk in men and women. Osteoporos Int 18:1109–1117.
4. Nguyen ND, Frost SA, Centre JR, et al. 2008 Development of prognostic nomograms for individualizing 5-year and 10-year fracture risks. Osteoporos Int 19:1431–1444.
5. Hippisley-Cox J, Coupland C. 2009 Predicting risk of osteoporotic fracture in men and women in England and Wales: prospective derivation and validation of QfractureScores. BMJ 339:b4229.
6. Lorenc R, Głuszko P, Karczmarewicz E, et al. 2013 Recommendations for diagnostic and therapeutic management in osteoporosis. Med Prakt 1–48, [in Polish].
7. Camacho PM, Petak SM, Binkley N, et al. 2016 American Association of Clinical Endocrinologists and American College of Endocrinology Clinical Practice guidelines for the diagnosis and treatment of postmenopausal osteoporosis—2016. Endocr Pract 22:1–42.
8. Włodarek D, Głąbska D, Kołota A, et al. 2014 Calcium intake and osteoporosis: the influence of calcium intake from dairy products on hip bone mineral density and fracture incidence—a population-based study in women over 55 years of age. Public Health Nutr 17:383–389.
9. Głąbska D, Włodarek D, Kołota A, et al. 2016 Assessment of mineral intake in the diets of Polish postmenopausal women in relation to their BMI-the RAC-OST-POL study: mineral intake in relation to BMI. J Health Popul Nutr 35(1):23.
10. Drozdzowska B, Wiktor K, Pluskiewicz W. 2013 Functional status and prevalence of falls and fractures in population-based sample of postmenopausal women from the RAC-OST-POL Study. Int J Clin Pract 67:673–681.
11. Pluskiewicz W, Adamczyk P, Czekajło A, et al. 2014 Influence of education, marital status, occupation, and the place of living on skeletal status, fracture prevalence, and the course and effectiveness of osteoporotic therapy in women in the RAC-OST-POL Study. J Bone Miner Metab 32:89–95.
12. Rokicki W, Drozdzowska B, Czekajło A, et al. 2014 Common ophthalmic problems of urban and rural postmenopausal women in a population sample of Raciborz district, a RAC-OST-POL Study. Ann Agric Environ Med 21(1):70–74.
13. Pluskiewicz W, Adamczyk P, Czekajło A, et al. 2012 Epidemiological data on osteoporosis in women from the RAC-OST-POL study. J Clin Densitom 15:308–314.
14. Pluskiewicz W, Adamczyk P, Czekajło A, et al. 2015 High fracture probability predicts fractures in a 4-year follow-up in women from the RAC-OST-POL study. Osteoporos Int 26:2811–2820.
15. WEKA download page. Available at: http://www.cs.waikato.ac.nz/ml/weka/downloading.html.
16. R Project for Statistical Computing, web page. Available at: https://www.r-project.org/.
17. Werner A, Bach M, Pluskiewicz W. 2016 The study of preprocessing methods' utility in analysis of multidimensional and highly imbalanced medical data. In: Proceedings of 11th International Conference IIIS2016. ISBN 978-83-65621-00-9.
18. PQStat—Precision and Quality of Statistical Analysis package, web page. Available at: http://pqstat.pl/.
19. Cramer JS. 1999 Predictive performance of the binary logit model in unbalanced samples. Statistician 48.
20. Collins GS, Mallett S, Altman DG. 2011 Predicting risk of osteoporotic and hip fracture in the United Kingdom: prospective independent and external validation of QfractureScores. BMJ 342:1–12.
21. Bach M, Werner A, Żywiec J, Pluskiewicz W. 2016 The study of under- and over-sampling methods' utility in analysis of highly imbalanced data on osteoporosis. Inf Sci (Ny) doi:10.1016/j.ins.2016.09.038.