New Approaches in Sport Sciences (NASS), Vol 2, No 4, December 2020, 87-100

Prediction of Osteoporosis by K- NN Algorithm and Prescribing Physical Activity for Elderly Women

Bakhtyan Tartibian
Professor of Exercise Physiology, Department of Exercise Physiology, Faculty of Physical Education and Sport Sciences, Allameh Tabataba'i University, Tehran, Iran
Leila Fasihi
M.Sc. Student, Department of Physical Education and sport science, Faculty of Physical Education and sport science, University of Allameh Tabatabai, Tehran, Iran
Rasul Eslami
Associate Professor of Exercise Physiology, Department of Exercise Physiology, Faculty of Physical Education and Sport Sciences, Allameh Tabataba'i University, Tehran, Iran

Received: Feb 17, 2021; **Accepted:** Feb 23, 2021

doi: 10.22054/nass.2021.59130.1088

Abstract

Purpose: A number of clinical decision tools for osteoporosis risk assessment have been developed to select postmenopausal women for the measurement of bone mineral density. We developed Data mining algorithm with the aim of more accurately identifying the risk of in postmenopausal women compared to the ability of conventional clinical decision tools. **Materials and Methods:** The present study was a cross-sectional development study conducted in the second half of 2018. In the present study, first, by identifying the influential variables, a survey questionnaire was prepared to select the most important clinical factors. Bone mineral density information of women referred to the bone density measurement unit of Khatam Al-Anbia Hospital in Tehran was used to teach the K-Nearest Ne neighbors (K-NN) algorithm (based on simple studies). Evaluation was based on accuracy. We also reviewed the results of several scientific articles and suggested the best sports activities according to the bone density of individuals. **Results:** The K-NN algorithm with sub-curve surface (AUC) showed significant performance. The algorithm predicted the risk of osteoporosis with an accuracy of 61.7% in the femoral neck for women participating in the experiment. Also, regular resistance and endurance training exercises repeated for 2-3 times a week for a year can have significant effects on maintaining or increasing hip BMD in postmenopausal women. **Conclusion:** Considering various pre¬dictors associated with low bone density, the K-NN algorithm may be an ef¬fective tool for identifying women at high risk for osteoporosis. This method widely recommends and predicts regular resistance and endurance training exercises for women with a high risk of osteoporosis.

***Keywords**:* Data Mining, Bone Mineral Density, Exercise, Aging

Author's e-mails: ba.tartibian@gmail.com, fasihi73@gmail.com **(Corresponding Author)**, eslami.rasul@gmail.com

INTRODUCTION

Osteoporosis is a common bone disease that is known to increase the risk of bone fractures in order to reduce bone density and loss of bone microstructure (Nayak et al, 2011). As a result of this, bone strength decreases and the risk of fractures increases (Iliou et al, 2017) In osteoporosis, bone mineral density (BMD) decreases and the microstructure of the bone is disrupted. Also, the concentration and types of proteins in the bones are altered. The most common osteoporotic fractures are fractures of the hip, vertebrae, and wrist, occurring in areas associated with low BMD, but the incidence increases after the age of 50 (Johnell & Kanis, 2006). According to the World Health Organiza¬tion, annually more than 9.8 million breakages occur due to osteoporosis (Sugimoto et al, 2016). It is estimated that osteoporosis will affect 200 million women worldwide (almost 1 out of 10 women aged over 60 years, 1 out of 5 women aged over 70 years, 2 out of 5 women aged over 80 years, and 2 out of 3 women aged over 90 years) (Johnell & Kanis, 2006). A recent study carried out by the Ministry of Health and Medical Education in Iran showed that osteoporosis affected 1 out of 4 Iranian women aged over 50 years old (Rohollahi, 2007). Often the first apparent symptom of osteoporosis is a broken bone. This is why the condition is also known as "the silent crippler", as many people do not realize that they have osteoporosis until it's too late (Iliou et al, 2017).

Osteoporosis is one of the major causes of disability and mortality in older people (Fredman et al, 2019). The mortality rate of hip fracture (femoral head) in the first year after the fracture occurred is about 20% in older people and half of these people will have some degree of disability for the rest of their lives. It is anticipated that more than 75% of osteoporosis fractures may occur in developing countries over the next 50 years (Sambrook et al, 1994). Iran, as well as other developing countries, will have a significant population of older people in the next 50 years (Rohollahi, 2007). Which demonstrates and highlights the importance of predicting and early detection of osteoporosis with available, simple and inexpensive tools (Rohollahi, 2007). Therefore, it seems that planning to predict and prevent osteoporosis should be one of the health priorities in Iran. Recently, some data mining methods can be useful in diagnosing the disease. Data mining, a set of analytical techniques for discovering, classifying and predicting, has the potential

to provide such insight. Compared to traditional statistical approaches, it offers several advantages for identifying intervention goals and strategies. These include using algorithms to identify the strongest predictions among hundreds of variables simultaneously, creating categories and refined cutoff values for a large number of variables, and providing optimized metrics for custom intervention purposes (e.g., group Refined in age). In the present study, To investigate this potential, our study used data mining methods to predict osteopenia and osteoporosis in women. Physical activity (PA) is one of the most effective tools to counter age-related health conditions. Regular PA has preventive effects on cardiovascular conditions, type 2 diabetes, and degenerative diseases such as osteoporosis and osteoarthritis (Murtagh et al, 2010). Regular exercise is widely recommended as the most effective non-pharmacological method for improving and maintaining BMD and can also reduce the risk of falling (Arazi et al, 2016). So far, a lot of research has been done on the amount of intervention in different sports activities that can cause a significant change in bone density and bone regeneration in postmenopausal women.

American College of Sports Medicine (ACSM) declare that individuals with purpose of bone health should have an active life style, so that aerobic exercises and activities (weight-bearing exercises) with intensity of moderate-to-high are recommended to maintain or to ascend bone mass (Bloomfield, 2004). Exercise increases bone mineral density, bone mass, bone strength and bone mechanical properties. It seems to directly or indirectly act on almost all the bone cell types and affect many aspects of bone remodeling (Yuan et al, 2016). The data from Kohrt et al (1997), reported that an exercise program including walking, jogging and stair climbing resulted in significant increases in BMD of the whole body, lumbar spine, femoral neck and Ward's triangle. However, high intensity exercises (ultra-marathon, running >64 km per week) is accompanied with a lot of osteoarticular damages and micro-injuries (Kohrt et al, 1997). In addition, the optimal level of physical activity promoting BMD benefits and modulating osteoprotectors is still unknown (Leme & do Carmo Sitta, 2013).

Generally, resistance exercise training has been identified as the most effective type of exercise on bone health. It is clear that a diverse loads or resistances are imposed to the bone during resistance exercise

training which induce stimuli and cause osteogenic response of the bone (Turner & Robling, 2005). Really, it appears that mechanical loading forces have lower effect in emerging an osteogenic effect with aging. Based this, likely a progressive loss of bone sensitivity happen related to chemical and physical signals (Rubin et al, 1992). Overall, data from the present study have shown that resistance exercise increases the BMD at the trochanter and total hip, balance and strength and that these effects are more influential than after endurance training in elderly females (Marques et al, 2011). Despite the results of the above studies, there is confusion in determining exercise appropriate for the condition of older women in the three groups of healthy, osteopenia and osteoporosis. There for the existence of a simple tool to suggest appropriate exercise activity to these three groups is essential. According to many authoritative scientific articles, appropriate physical activity can be suggested to women. Therefore our first purpose is to suggest the best and most effective sports activities to the three categories of healthy women, osteopenia, and osteoporosis using the article's results. The second purpose of this study was to suggest appropriate exercise by reviewing the exercises performed to improve bone density in valid scientific articles.

METHOD

To do effective data mining., in addition to the need for relevant information, the appropriate data mining method must also be used. The appropriate data mining method includes all the steps of data mining such as data collection, data preparation, modeling and evaluation required.

Description and Data collection

The clinical information used in this study was related to 327 patients referred to Khatam Al-Anbia Hospital in Tehran in 2018. Records were related to lifestyle, personal information, and illness information. The study was approved by local institutional review committees, and participants signed informed consent forms. Inclusion criteria were women between 65 and 85 years old, with medical records and clinical trials in hospital, T-score between (-3/5< T-score <2, BMI between 18 and 40, accessibility by phone or internet. Exclusion criteria included people who were being treated for osteoporosis or estrogen hormones or

had a history of taking these drugs, or had conditions such as chronic illness.

Data modeling

K-NN algorithm is a supervised machine learning algorithm that can be used to solve both classification and regression problems. Such a method is a classification method in data mining that is used specifically for large data sets. A set that has many predictions, such as the ones used in bioinformatics, is called a large data set. This method is used when a lot of data such as medical data sets are available (Huang et al, 2018). In this method, the data were first divided into two groups: training (for training) and test (for evaluation). Data classification was done completely randomly using Partition node in Clementine software training data in which training data had 75% of the data and test data had the remaining 25% of the data (Koski & Noble, 2011). (figure 1)

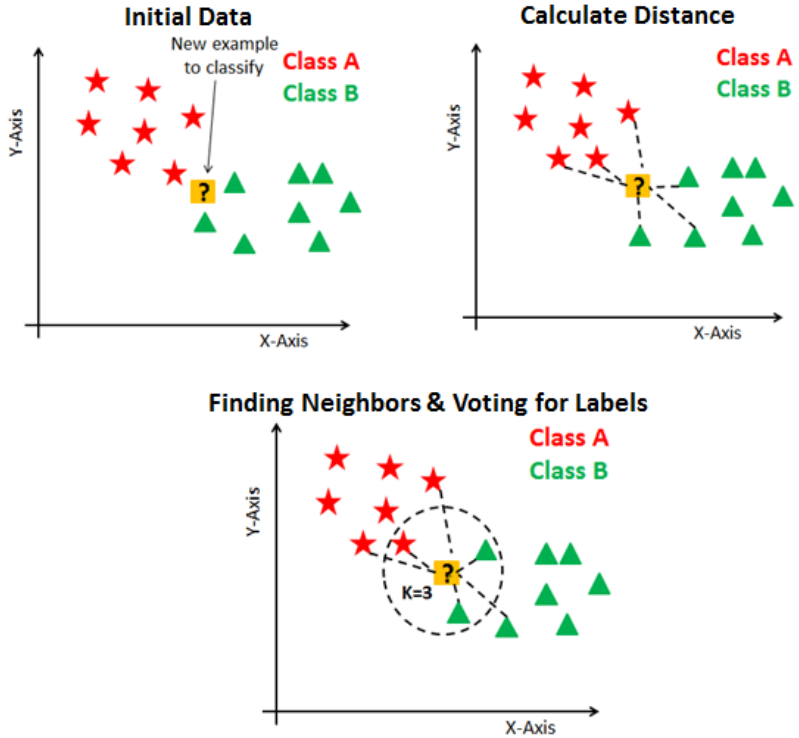

Figure 1

K-NN algorithm was selected in the Clementine software. Primarily, 9 features were entered in the elementary configuration of this model. These 9 features have been selected using dimensional reduction methods (K-NN algorithm and the opinion of specialists. An K-NN model was run using these 9 features and its results were obtained.

Clearing and normalizing data

Data is usually not suitable for processing immediately after collection. To make the best use of data, it must be modified to fit data mining algorithms. Zero and one were used for the questions that had yes or no answers. Zero meant a no answer and number one meant a yes answer. At this stage, the data set in Excel format was transferred to MATLAB software version 2020 and analyzed.

Separation criteria according to (T-Score) for K-NN prediction

In the case of data mining to diagnose osteoporosis, the cleansing and normalization phase was performed. The data set was imported into an Excel file to be used to teach algorithms. In this study from 9 variables, an input and a target (T-score) were considered women's data: T-score <-1 Healthy, T-score between -1 to -2.5 Osteopenia, and T- score> -2.5 was Osteoporosis.

RESULTS

The clinical history used in this study was 327 women with 80 healthy and with 247 patients suffering from osteoporosis and osteopenia. The following table provides a brief description of the data.

Table 1: Characteristics of the participants

Indicators	**Women**
	Mean and standard deviation
Age (years)	73. 70 ± 9.67
Height (cm)	159. 45 ± 6.23
Weight (kg)	72.92 ± 12.30
Body mass index (kg / cm2)	28.70 ± 4.78

In order to identify the necessary indicators for the diagnosis of osteoporosis, first by studying the relevant books and articles, clinical

factors related to osteoporosis were extracted. Then, by reviewing similar articles, and consulting with a specialist physician, and distributing and collecting relevant questionnaires from the subjects, the most important clinical factors of osteoporosis were prepared (Table 2).

Table 2: The risk factors (attributes) in our dataset

Number	Features
1	Age
2	Height
3	Weight
4	BMI
5	back pain
6	Family history
7	Vitamin D intake
8	physical activity
9	Calcium intake

K-Nearest Neighbors model: This model had the highest classification precision using 9 features, which seems to be an appropriate. According to the results of the confusion matrix of the K-nearest neighbor algorithm, it can be seen that this algorithm can predict and evaluate healthy women, osteopenia and osteoporosis with 61.7% accuracy. Figure (3)

Table 3:

Percentage error		65	65.2	46.7		
Correct percentage		35	34.8	53.3		
True class	Healthy	13	7	0	65	35
	Osteopenia	7	30	8	66.7	33.3
	Osteoporosis	0	9	7	43.8	56.3
		Healthy	Osteopenia	Osteoporosis	Correct percentage	Error percentage
		Predicted class				

Using the results of Table 4, it can be seen that the most effective sports activities in terms of age and T-score were recommended to healthy postmenopausal women, osteopenia and osteoporosis according to the results of valid scientific articles.

Table 4: Sports activities in terms of age and T-score

Authors	**Age**	**Exercise suggestion**	**Time**	**T-score**
Marques and Mota, 2011	55-85	muscle endurance exercises, 1–3 sets, 8–15 reps, 1RM (more details n.g.), 10 min balance & dynamic exercise (walking, playing with ball, rope, sticks, etc.), 10 min agility training (coordination, balance, ball games, dance)	Training was performed twice a week, 50 min	Healthy (T-score≥-1)
Englund et al.	55-85	muscle strength, balance, handgrip strength, walking performance and coordination exercises.	Training was performed 12-month, twice a week, 50 min	Osteopenia (-2/5< T-score <-1)
Liu, 2015	55-85	The exercise consisted of daily outdoor walking, the intensity of which was 50% of maximum oxygen consumption, with a duration of at least 1 h with more than 8000 steps, at a frequency of 4 days a week, over a 12-month period.	6Month 3 × daily ≈ 3–5, HE (96%)	Osteoporosis (T-score ≤ -2/5)

DISCUSSION

Based on various data mining techniques, we explored a new approach to predicting the risk of osteoporosis in postmenopausal women using data from several hospitals. The results of the present study showed that

the K-NN algorithm was more accurate in predicting osteoporosis for the three groups of healthy, osteopenia and osteoporosis. Another result of the present study was to suggest effective physical activity to these three groups to increase BMD.

Among the data mining techniques and conventional methods K-NN algorithm discriminated more accurately between Old menopausal women. K-NN algorithm was more effective in analyzing the epidemiological underlying patterns of osteoporosis compared with the other methods. This finding is consistent with a previous study on the com¬parison of data mining techniques in various complex discriminating problems for predicting disease (Agranoff et al, 2006). Most ex¬perts have used conventional methods, including OST, ORAI, SCORE, and OSIRIS, because of their simplicity (Schwartz & Steinberg, 2006). Our method is useful because computerized diagnostic deci¬sion supports have been increasingly easy to access due to the advancement of information systems for many medical problems (Patel et al, 2008).

If our prediction model retains good performance after validation in a larger population, it will be possible to use this technique as a cost-effective prescreening tool to deter¬mine candidates for evaluation with DEXA and also to pre¬vent osteoporotic fracture in postmenopausal women at high risk. The patients in the high risk group categorized by this method should receive DEXA screening at the hospital. However, patients in the low risk group could postpone re¬ceiving a DEXA scan. Women experience menopause at 50 years old on average (Yeun, 2000).

There are several limitations to this study. First, the study was based on a cross-sectional survey which has several defects according to a medical view. For example, the prev¬alence of disease was based on a health interview survey taken on one occasion. Weight, height, body mass index, and hormone therapy status, as well as BMD, could differ according to time of the measurement. Second, it was diffi¬cult to consider drug effects. For example, treatment with steroids for rheumatoid arthritis, asthma, dermatitis, or auto¬immune diseases is known to cause glucocorticoid-induced osteoporosis (Van Staa et al, 2005). Systematic approaches are warranted to con¬sider the long-term effects of drugs in further studies. Third, our study was charac¬terized by imbalanced class distribution. Traditional classi¬fiers have generally shown poor performance on imbal¬anced data sets because they are

designed to find the best classification for the majority (Sun et al, 2007). The imbalanced class was still critical even though we adopted a dense grid search to decide the optimal prediction models in order to overcome this problem. Therefore, if more patient data associated with osteoporosis were collected, the performance would be ex¬pected to improve.

Another result of the present study was to suggest effective physical activity to these three groups to increase BMD. Regular exercise is widely recommended as the most effective non-pharmacological method for improving and maintaining BMD and can also reduce the risk of falling (Arazi et al, 2016). This finding is consistent with several studies. The data from Kohrt et al (2011). reported that an exercise program including walking, jogging and stair climbing resulted in significant increases in BMD of the whole body, lumbar spine, femoral neck (Kohrt et al, 1997). Zehnacker et al. has declared that performing resistance exercise training regularly during one year with frequency of 2-3 times per week has considerable effects to maintain or increase the BMD of spine and hip in postmenopausal female (Zehnacker & Bemis-Dougherty, 2007). Maud et al (2006) designed the study of a long-distance runner older than 70 years who had training history more than 50 years. But, they did not find any alterations in musculoskeletal system that, it was not consistent with the results of the present study (Maud & ML, 1981). Overall, data from the second part of the present study have shown that Regular Exercise increases the BMD at the femoral neck.

CONCLUSIONS

The most important finding of this study is the identification of postmenopausal women at high risk of osteoporosis to increase the possibility of appropriate treat¬ment before fracture occurs. Machine learning methods might contribute to the advancement of clinical decision tools and understanding about the risk factors for osteopo¬rosis. Further studies should be targeted at constructing an extended prediction model for progressive osteoporosis through the collection of prospective data, and the simulta¬neous prediction of osteopenia and osteoporosis using multi-category classification. We hope that this study enables women to reduce the risk of osteoporosis. Also, regular resistance and endurance training exercises repeated for 2-3 times a week for a year can have significant effects on maintaining or increasing hip BMD in postmenopausal women.

REFERENCES

Agranoff, D., Fernandez-Reyes, D., Papadopoulos, M. C., Rojas, S. A., Herbster, M., Loosemore, A., . . . Pollok, R. (2006). Identification of diagnostic markers for tuberculosis by proteomic fingerprinting of serum. *The Lancet Journal, 368*(9540), 1012-1021.

Arazi, H., Eghbali, E., Saeedi, T., & Moghadam, R. (2016). The relationship of physical activity and anthropometric and physiological characteristics to bone mineral density in postmenopausal women. *Journal of Clinical Densitometry, 19*(3), 382-388.

Fredman, L., Ranker, L. R., Strunin, L., Smith, M. L., & Applebaum, K. M. (2019). Caregiving intensity and mortality in older women, accounting for time-varying and lagged caregiver status: The Caregiver-Study of Osteoporotic Fractures study. *The Gerontologist Journal, 59*(5), e461-e469.

Huang, S., Cai, N., Pacheco, P. P., Narrandes, S., Wang, Y., & Xu, W. (2018). Applications of support vector machine (SVM) learning in cancer genomics. *Cancer Genomics-Proteomics Journal, 15*(1), 41-51.

Iliou, T., Anagnostopoulos, C.-N., Stephanakis, I. M., & Anastassopoulos, G. (2017). A novel data preprocessing method for boosting neural network performance: a case study in osteoporosis prediction. *Information Sciences Journal, 380*, 92-100.

Johnell, O., & Kanis, J. (2006). An estimate of the worldwide prevalence and disability associated with osteoporotic fractures. *Osteoporosis International Journal, 17*(12), 1726-1733.

Kohrt, W. M., Ehsani, A. A., & Birge JR, S. J. (1997). Effects of exercise involving predominantly either joint-reaction or ground-reaction forces on bone mineral density in older women. *Journal of Bone and Mineral Research, 12*(8), 1253-1261.

Koski, T., & Noble, J. (2011). *Bayesian networks: an introduction* (Vol. 924): John Wiley & Sons. *journal of network and computer applications, 12*(9), 44-66.

Leme, L. E. G., & do Carmo Sitta, M. (2013). Bone Mineral Density and High-Performance Aerobic Activity in Older Adults Experience in Brazil. *Topics in Osteoporosis Journal*, 193-205.

Marques, E. A., Wanderley, F., Machado, L., Sousa, F., Viana, J. L., Moreira-Gonçalves, D., . . . Carvalho, J. (2011). Effects of resistance and aerobic exercise on physical function, bone mineral density, OPG and RANKL in older women. *Experimental gerontology, 46*(7), 524-532.

Maud, D., & ML, F. (1981). Fifty years of training and competition in the marathon: Wally Hayward, age 70-a physiological profile. *South African Medical Journal*, *59*(5), 153-157.

Murtagh, E. M., Murphy, M. H., & Boone-Heinonen, J. (2010). Walking–the first steps in cardiovascular disease prevention. *Current opinion in cardiology*, *25*(5), 490.

Nayak, S., Roberts, M. S., & Greenspan, S. L. (2011). Cost-effectiveness of different screening strategies for osteoporosis in postmenopausal women. *Annals of Internal Medicine*, *155*(11), 751-761.

Patel, V. L., Shortliffe, E. H., Stefanelli, M., Szolovits, P., Berthold, M. R., Bellazzi, R., & Abu-Hanna, A. (2008). coming of age of artificial intelligence in. *journal of network and computer applications*, *27*(1), 33-46.

Rohollahi, F. (2007). Take a look at prevalence of osteoporosis in the world and Iran. *Hamshahri newspaper. Health part*, *4150*.

Rubin, C. T., Bain, S. D., & McLeod, K. J. (1992). Suppression of the osteogenic response in the aging skeleton. *Calcified tissue international*, *50*(4), 306-313.

Sambrook, P., Dequiker, J., & Rasp, H. (1994). metabolic bone disease, Report of a WHO study group, Assessment of fracture risk and its application to screening for postmenopausal Osteoporosis. *WHO Technical Report Series, Geneva*, *5*(3), 124-137.

Schwartz, E. N., & Steinberg, D. M. (2006). Prescreening tools to determine who needs DXA. *Current osteoporosis reports*, *4*(4), 148-152.

Sugimoto, T., Sato, M., Dehle, F. C., Brnabic, A. J., Weston, A., & Burge, R. (2016). Lifestyle-related metabolic disorders, osteoporosis, and fracture risk in Asia: A systematic review. *Value in health regional issues*, *9*, 49-56.

Sun, Y., Kamel, M. S., Wong, A. K., & Wang, Y. (2007). Cost-sensitive boosting for classification of imbalanced data. *Pattern Recognition*, *40*(12), 3358-3378.

Turner, C. H., & Robling, A. G. (2005). Mechanisms by which exercise improves bone strength. *Journal of bone and mineral metabolism*, *23*(1), 16-22.

Van Staa, T., Leufkens, H., Abenhaim, L., Zhang, B., & Cooper, C. (2005). Use of oral corticosteroids and risk of fractures. In: Wiley Online Library. *Journal of bone and osteoporosis*, *25*(2), 19-30.

Yeun, E. (2000). A study on the health promoting lifestyle practices of middle-aged women in Korea. *J Korean Soc Health Educ Promot*, *17*(1), 41-59.

Yuan, Y., Chen, X., Zhang, L., Wu, J., Guo, J., Zou, D., . . . Zou, J. (2016). The roles of exercise in bone remodeling and in prevention and treatment of osteoporosis. *Progress in Biophysics and Molecular Biology, 122*(2), 122-130.

Zehnacker, C. H., & Bemis-Dougherty, A. (2007). Effect of weighted exercises on bone mineral density in post menopausal women a systematic review. *Journal of geriatric physical therapy, 30*(2), 79-88.