Journal of
Cardio - Thoracic Medicine

http://jctm.mums.ac.ir

Determinants of Serum Vitamin D level; A Data Mining Approach

Zahra Amiri [1,2*], Fahimeh Moafian [1,2*], Payam Sharifan [2,3,4*], Elahe Hasanzadeh [4*], Zahra Khorasanchi [3], Moniba Bijari [2], Maryam Mohammadi Bajgiran [2], Sara Saffar Soflaei [2], Susan Darroudi [2], Hamideh Ghazizadeh [2,4], Fatemeh Rouhi [4], Fatemeh Mohseni Nik [4], Maryam Tayefi [5], Samira Roohi [4], Ayad Noor [6], Abdullah Al yakobi [6], Mohammed Hadi Lafta Alboresha [6], Gordon A. Ferns [7], Habibollah Esmaily [8**], Majid Ghayour-Mobarhan [2,3**]

[1] Department of Pure Mathematics, Center of Excellence in Analysis on Algebraic Structures (CEAAS), Ferdowsi University of Mashhad, P.O. Box 1159, Mashhad 91775, Iran.
[2] International UNESCO Center for Health-Related Basic Sciences and Human Nutrition, Mashhad University of Medical Sciences, Mashhad, Iran.
[3] Department of Nutrition, School of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran.
[4] Student Research Committee, School of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran.
[5] Norwegian Center for E-Health Research, University Hospital of North Norway, Tromsø, Norway.
[6] Department of Biology, Faculty of Sciences, Mashhad Branch, Islamic Azad University, Mashhad, Iran.
[7] Division of Medical Education, Brighton & Sussex Medical School, Falmer, Brighton, Sussex, UK.
[8] Social Determinants of Health Research Center, Mashhad University of Medical Sciences, Mashhad, Iran.

ARTICLE INFO

Article type:
Original Article

Article history:
Received: 25 May 2023
Revised: 2 July 2023
Accepted: 18 July 2023

Keywords:
Vitamin D
Oxidative Stress
Machine Learning
Data Mining Algorithm

ABSTRACT

Introduction: Serum vitamin D levels are related to a wide spectrum of factors including low sunlight exposure, high oxidative stress, low physical activity and sleep disorders. In this paper we are going to investigate the most crucial parameters associated with serum vitamin D levels in survey of ultraviolet intake by nutritional approach (SUVINA) study with a data mining approach.

Material and Methods: Data including demographic, anthropometric, clinical and laboratory information were extracted from the SUVINA dataset comprising 289 subjects who were enrolled into our study. The XGBoost algorithm was used to define the most important features related to vitamin D level in our population.

Results: Applying XGBoost modeling for vitamin D level showed that the presented scheme can determine the most important determinants of serum vitamin D level with an accuracy of 91%. Pro-oxidant anti-oxidant balance (PAB), body fat percentage, physical activity level (PAL), age, restless leg syndrome (RLS), and dietary inflammatory index (DII) density were the most important variables correlated with vitamin D deficiency.

Conclusion: Using XGBoost and with an accuracy of more than 90%, we showed that the six most important risk factors for vitamin D deficiency are PAB, PAL, age, body fat percentage, RLS and DII density, respectively.

► *Amiri, Z., Moafian , F., Sharifan, P., Hasanzadeh, E., Khorasanchi, Z., Bijari, M., Mohammadi Bajgiran , M., Saffar Soflaei, S., Darroudi, S., Ghazizadeh, H., Rouhi, F., Mohseni Nik, F., Tayefi , M., Roohi, S., Noor, A., Al yakobi, A., Lafta Alboresha, M.H., Ferns, G.A., Esmaily, H., Ghayour-Mobarhan, M. Determinants of Serum Vitamin D level; A Data Mining Approach. J Cardiothorac Med. 2023; 11(3): 1207-1215. Doi: 10.22038/jctm.2023.72615.1423*

* Equal as a first author
** Corresponding Authors: 1. Majid Ghayour-Mobarhan, MD, Ph.D; Metabolic Syndrome Research Center, Mashhad University of Medical Sciences, Mashhad, Iran. Postal code: 99199-91766. Tel: +98 513 800 2288, Fax: +98 513 800 2287. Email: ghayourm@mums.ac.ir .
2. Habibollah Esmaily, Ph.D: International UNESCO center for Health-Related Basic Sciences and Human Nutrition, Mashhad University of Medical Sciences, Mashhad, Iran. Postal code: 99199-91765. Tel: +98 513 800 2271, Fax: +98 513 800 2287. Email: esmailyh@mums.ac.ir .

Introduction

Vitamin D deficiency is one of the most prevalent micronutrient deficiencies (1, 2). It is estimated that more than one billion people are affected with vitamin D deficiency which makes it a significant global concern (1, 3, 4). Approximately 80% of Iranian adolescents are vitamin D deficient and the prevalence is even higher among elderly (5, 6). Vitamin D is crucial to the health, survival and fertility of human beings and its deficiency has been connected to many diseases such as many cancers and even obesity (1, 3, 7). Hence, it is of great importance in the management of this situation to not only identify risk factors for vitamin D deficiency but also to determine to which degrees these factors are effective.

Despite the fact that low sunlight exposure is the primary known risk factor for vitamin D deficiency, recent studies have determined several other factors influencing vitamin D levels (8, 9). However, existing medical literature in this filed is rather non-conclusive. While physical activity level (PAL) has been shown to boost serum vitamin D levels in general population (10), some studies did not find significant difference in vitamin D levels between active and sedentary participants (11)]. Other issues including age-related changes in vitamin D level (12, 13), oxidative stress markers and their relation with vitamin D deficiency (5) and the impact of body fat percentage on the vitamin D level (14-17) were also the subject of recent studies. However, due to the small sample size, observational nature of most of the studies in this field, and confounding variables, the results are rather contradictory and inadequate.

Data mining and machine learning are promising estate-of-the-art techniques dealing with big datasets and have been adopted in medical research with great success (18-23). In recent medical research these techniques are used especially for developing predictive models (24-28). In this article, we have used Xtreme Gradiant Boosting (XGBoosting) as the model, because of its high performance in the context of vitamin D deficiency risk factors (29).

The aim of the present paper was to contribute to a deeper understanding of vitamin D deficiency risk factors and their effect using machine learning and decision-tree models.

Materials and Methods

Clinical Trial

This paper is based on a total-blind clinical trial, applying fortified low-fat dairy products including nano-encapsulated vitamin D (1500 IU). It was done between January 2019 and March 2019 in Mashhad, Iran (for 2.5-month). The protocol was approved by National Institute for Medical Research Development (NIMAD) (protocol ID: IR.NIMAD.REC.1396.027) before starting the study. This paper was a pilot study (Trial registration: IRCT20101130005280N27, wwww.IRCT.ir).

Study Population

This cross-sectional research is a part of the SUVINA study which is discussed in details elsewhere (30). In summary, the participants were recruited from Mashhad University of Medical Sciences staffs and students from which 289 met the eligibility criteria and finished the trial (146 female, 143 male). The inclusion criteria were as following: subjects who were 30 and 50 years old, with central obesity, defined according to the International Diabetes Federation (IDF) criteria (31), reported energy intake (800-4200 kcal/d), no history of chronic diseases such as liver, renal, malignancy, bone disorders and thyroid, and as well as no drug history of agents interacting with bone metabolism and vitamin D (such as diuretics, corticosteroids, and bisphosphonates), vitamin D supplements, and medications for mood or sleep disorders.

Individuals who completed less than 10% of the food frequency questionnaire (FFQ) were excluded (6). Also, we excluded those with weight changes more than 5 Kg over the last year or those with especial diets such as vegetarianism and participants who smoke or consume alcohol. Additionally, those with missing data regarding metabolic syndrome (MetS) components (waist circumference, systolic and diastolic blood pressure, serum blood glucose, triglyceride, and high-density lipoprotein) were also excluded. For all the participants informed written consent was obtained. The research ethics committee of NIMAD approved the study (protocol ID: IR.NIMAD.REC.1396.027).

Measurements

Data regarding age, gender, marital status, education level, smoking status and PAL were obtained during an interview with healthcare professionals using validated questionnaires described previously (30, 32, 33). Five certified nutritionists from Mashhad University of Medical Science approved the 65 dietary items of FFQ that is applied to evaluate dietary intake and the patients were divided into low and high antioxidant-dietary patterns based on the questionnaire. Also dietary inflammatory index (DII) and the healthy eating index (HEI), were evaluated according to validated FFQ (34). Other validated questionnaires such as depression anxiety stress scale (DASS-42), Pittsburg sleep quality index (PSQI), restless legs syndrome (RLS), Epworth, insomnia, premenstrual symptoms screening tool (PSST), apnea and quality of life (QoL) were used as described previously (30).

Anthropometric variables including weight, height, body mass index (BMI) and the circumference of hip, waist, neck and mid-arm were also assessed according to standard protocols (35). We also used a bio-impedance device (TANITA BC 418, Japan) to determine total body water (TBW), Body fat mass, body fat percentage, free fat mass (FFM), trunk fat and basal metabolic rate (BMR). Other variables such as systolic blood pressure (SBP), diastolic blood pressure (DBP), pro-oxidant-oxidant balance (PAB) was determined according to the scheme determined in (5). MetS and hypertension were measured based on the IDF criteria (31) and diabetes mellitus were measured based on the scheme in (36).

Biochemical and hormonal features including serum 25 (OH) vitamin D, complete cell count (CBC) parameters, lipid profile, fasting blood sugar (FBS), iron, magnesium, calcium, creatinine, phosphorus, blood urine nitrogen (BUN), liver enzymes and bilirubin, creatine phosphokinase (CPK), uric acid, total protein and albumin, high sensitivity C-reactive protein (Hs-CRP), serum insulin, homeostatic model assessment for insulin resistance index (HOMA-IR), quantitative insulin sensitivity check index (QUICKI) were also measured based on protocols described previously (30, 37).

Data Analysis

One of the reliable and powerful algorithms is XGBoost that provides applicable results for various complex problems. It is an implementation of gradient boosted decision trees designed for speed and performance.

Sometimes it happens that we should select the important features for training data when there is a large number of features, because it takes much computational cost to train the dataset. This importance is calculated explicitly for each attribute in the dataset, allowing attributes to be compared to each other, giving an importance score for each attribute. Importance scores display how useful or valuable each feature was within the model, which means the more an attribute is used, the higher its relative importance.

An advantage of using gradient boosting is that it is simple to obtain importance scores for each feature. Importance is calculated for a single decision tree by the amount that each attribute split point improves the performance measure, weighted by the number of observations the node is responsible for. There are two properties that could be considered in any appropriate feature attribution method:

1. Accuracy. The sum of all the feature importance should sum up to the total importance of the model.
2. Consistency. The attributed importance of that feature should not diminish when we modified a model that relies more on a feature.

Here we choose the "weight" criterion to select feature importance of our dataset as among all criteria it performs well for both accuracy and consistency. Weight criterion scores is measured according to the number of times a feature is used to split the data across all trees. After selecting important features, we use XGBoost classifier model to fit it to testing and training subsets. In order to build more robust model, we need to consider different parameters and their values to be specified. This model requires parameter tuning to improve and fully leverage its advantages over other algorithms. Python version 3.8.6 was applied to interpret our data using XGBoost classifier.

Here we used the XGBoost model on 77 features such as: vitamin D, age, gender, educational level, weight, height, BMI, mid-arm circumference (MAC), wrist circumference, neck circumference, waist circumference (WC), hip circumference (HC), DBP, SBP, PAL, PAB, RLS, QoL, apnea, insomnia, Epworth, PSIQ, PSST, depression, anxiety and stress score, high/low antioxidant dietary pattern, DII density, HEI, MetS, hypertension, diabetes, TBW, FFM, BMR, trunk fat, body fat mass and percentage, platelets, red and white blood cell count, red cell distribution width, hematocrit, hemoglobin, mean corpuscular hemoglobin concentration, mean corpuscular hemoglobin, mean corpuscular volume, mean platelet volume, triglyceride, total cholesterol, FBG, high and low-density lipoprotein cholesterol, creatinine, BUN, calcium, phosphorus, iron, magnesium, Hs-CRP, total and direct bilirubin, alanine aminotransferase, alkaline phosphatase, aspartate aminotransferase, total protein, gamma-glutamyl transferase, and albumin, CPK, uric acid, serum insulin, HOMA-IR, QUICKI.

Results

Figure 1 shows the plot of the fifteen most important selected features including their scores with respect to the target. We set Y as "vitamin D" and the target and the other features were set as X. For selecting the most important variables, XGBoost classifier with weight measure was used and hyper-parameters were tuned as eta=0.01, col-sample-by-tree=1, subsample=0.6, and max-depth=5. We chose the 6 top variables with the important score and the greatest impact on the target as: PAB, PAL, age, body fat percentage, RLS, DII density. Figure 2 depicts the relative percentage of importance for the selected features and their attributed scores. In accordance with the relative percentage, PAB was the most important feature (36.21%) followed by PAL (19.23%), age (13.46%), body fat percentage (10.57%), RLS (10.25%) and DII density (10.25%). Using tertiles, we categorized vitamin D levels as the target into three classes, with less than 33% in class 0, 33% in class 1, and more than 33% in class 2.

We split X and Y into train and test subsets. A ratio of 85/15 was used for the data splitting where 85% were used for the train subset and 15% for the test subset. The XGBoost classifier model trained on the train dataset and then evaluated the model on the test dataset, such that the same feature selection approach, given us 91% accuracy.

Discussion

Vitamin D is an essential nutrient that has been linked to various diseases, cancers, and even overweight and obesity (7). To date, studies regarding vitamin D deficiency and its risk factors have had considerable limitations and have been inconsistent, and as we know, there is not an attempt to prioritize the importance of risk factors for vitamin D deficiency. According to our study, among the fifteen most important features for vitamin D level, the top six were PAB, PAL, age, body fat percentage, RLS and DII density, respectively.

An important finding of our study was that serum PAB had the greatest impact on serum vitamin D levels; this was twice that for PAL. These outcomes appear to be consistent with earlier studies, which showed that in patients with low levels of vitamin D, the activity of catalase, an antioxidant enzyme, is considerably decreased in comparison to healthy people (38). However, research on PAB indicated that high dose vitamin D supplementation increased the serum PAB levels in teenage girls. Although this study had some limitations including the lack of control group, short intervention period and high dose of vitamin D compared to previous studies, it highlights the need for further prospective and clinical trial studies to clarify the exact relationship (5).

There is a large body of evidence on the effect of PAL on vitamin D levels. Our observation is consistent with that of Ha et al. who found that in the Korean pediatric population, serum vitamin D levels were positively related to accelerometer-based PAL including low, moderate, and vigorous levels which is consistent with other studies on this population (10, 39, 40). Bauer et al. in a 2020 article stated that physical activity is

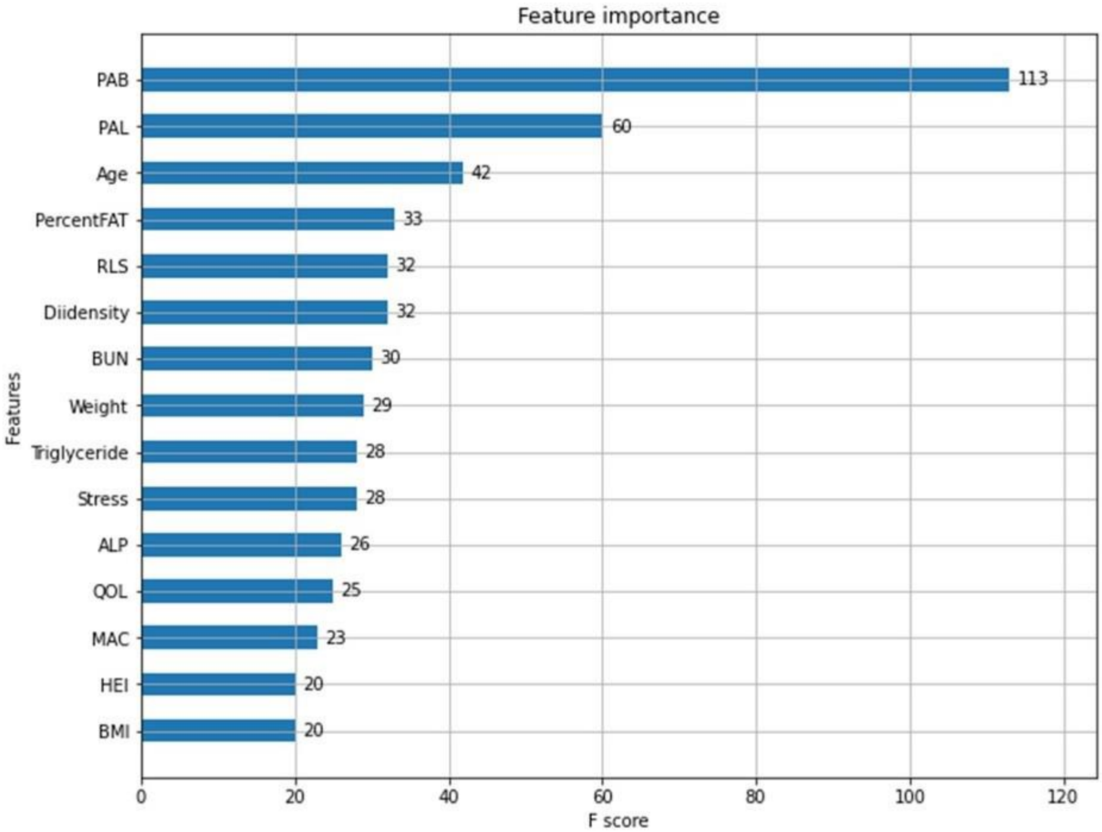

Figure 1. Vitamin D related factors; 15 most important features related to vitamin D. PAB pro-oxidant anti-oxidant balance, PAL physical activity level, RLS restless leg syndrome, DII density dietary inflammatory index density, BUN blood urea nitrogen, ALP alkaline phosphatase, QOL quality of life, MAC mid-arm circumference, HEI healthy eating index, BMI body mass index.

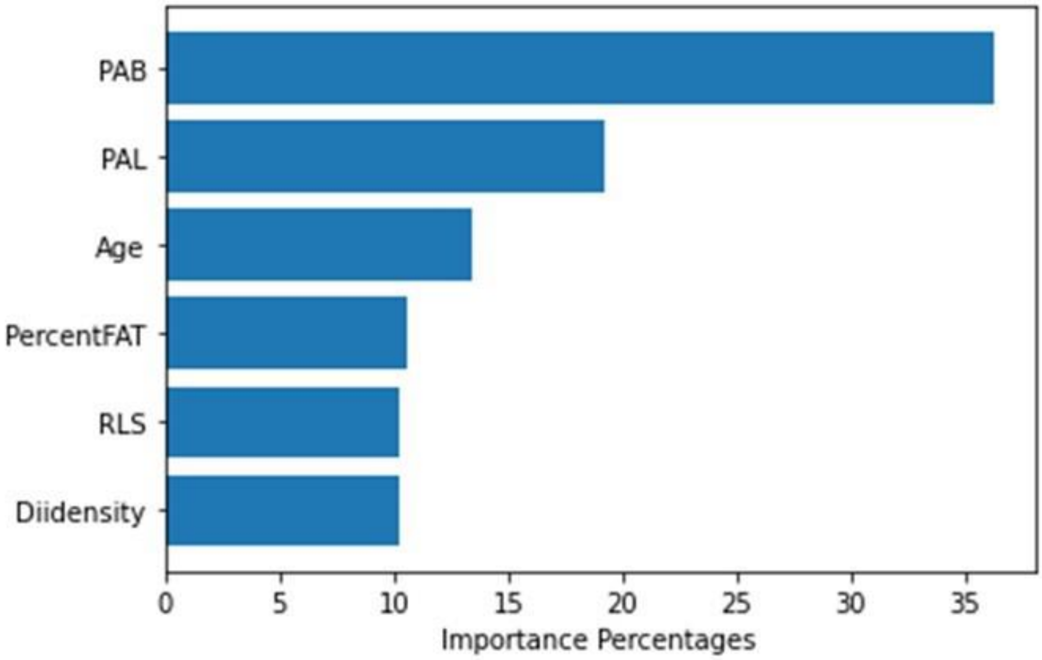

Figure 2. The importance percentage of the six most important features related to vitamin D. PAB pro-oxidant anti-oxidant balance, PAL physical activity level, RLS restless leg syndrome, DII density dietary inflammatory index density.

associated with higher concentrations of vitamin D independent of sunlight exposure duration in multiple sclerosis (MS) patients (41). However, some studies were unable to found a clear association between physical activity and serum vitamin D levels (11, 42). This controversy of findings could be due to the cross-sectional design of the researches, different methods of analysis and exclusion of frailest participants.

Findings of the present study were in line with previous research on the age-related changes in vitamin D levels and the effect of body fat percentage on this vitamin. Many articles showed that with aging, not only the cutaneous synthesis of vitamin D is decreased, but also the dietary consumption of vitamin D is declined which makes elderly more susceptible to vitamin D deficiency (43-45). Moreover, in a 2020 study, Leiu et al. observed that vitamin D insufficiency was more likely to happen in older women with high body fat percentage and low dairy product consumption (46). In the same vein, many researches such as that of Ong et al. reported the association between body fat percentage and anthropometric measures with vitamin D deficiency especially in the elderly (46-49). Conversely, Grønborg et al. reported no significant association between vitamin D and body fat percentage status (50). In this study, the sample size was small, which may raise some controversy.

We also found a RLS as a risk factor for vitamin D deficiency. These results are similar to those of Wali et al and Balaban et al. who found an inverse association between vitamin D levels and RLS (51, 52). A previous study has reported that vitamin D supplementation improves the severity of RLS symptoms (53). On the other hand, this concept has recently been challenged by Romano et al. demonstrating there is still limited evidence to substantiate the role of vitamin D supplementation in preventing or treating sleep disorders such as RLS (54). Another clinical trial was also unable to found improvement in RLS symptoms with vitamin D supplements (55). However, this study had some weaknesses which could affect the results such as short follow-up period, inadequate sample size, no objective measure of RLS symptoms and the loss of patients during follow-up.

Our analysis provides new insights into DII density and vitamin D relationship. As far as we know no study has examined the direct relationship between vitamin D and DII density factor so for the first time the present research explores this relationship using a decision tree.

Strengths and Limitations

To the best of our knowledge, this is the first study investigating factors associated with vitamin D status, using a machine learning method. The powerful and reliable algorithm of XGBoost provided advanced findings for many complex issues, particularly when accuracy and speed is needed. The limitations of the present study are the rather small sample size and observational design. Notwithstanding the limitations, the findings reported here shed new light on our understanding of factors affecting serum vitamin D level. Using the thriving technology of AI and machine learning to make a predictive model of factors with the greatest effect on the vitamin D deficiency, will help to prevent and reduce the prevalence of this disease.

Conclusion

Despite similarities and contradictions between the effects of different factors on Vitamin D, our results showed that PAB, PAL, age, body fat percentage, RLS, and DII density are the most important risk factors for vitamin D deficiency in an Iranian population and they are crucial to consider in order to reduce vitamin deficiency in our society. For the future works, we can consider the other data mining algorithms for analyzing the data.

Abbreviations

Physical Activity Level: **PAL** , Xtreme Gradient Boosting: **XGBoosting** , Survey of Ultraviolet Intake by Nutritional Approach : **SUVINA** ,Mashhad University of Medical Sciences: **MUMS** , International Diabetes Federation: **IDF**, Food Frequency Questionnaire: **FFQ** , National Institute for Medical Research Development: **NIMAD** , Dietary Inflammatory Index: **DII** , Healthy Eating Index: **HEI** , Depression Anxiety Stress Scale: **DASS** , Pittsburg Sleep Quality Index :

PSQI, Restless Legs Syndrome: **RLS**, Premenstrual Symptoms Screening Tool: **PSST** , Quality of Life: **QoL**, Body Mass Index: **BMI** , Total Body Water: **TBW**, Free Fat Mass: **FFM** , Basal Metabolic Rate: **BMR** , Systolic Blood Pressure: **SBP**, Diastolic Blood Pressure: **DBP**, Pro-oxidant Antioxidant Balance: **PAB**, Metabolic syndrome: **MetS**, Complete Cell Count: **CBC**, Fasting Blood Sugar : **FBS**, Blood Urine Nitrogen: **BUN**, High sensitivity C-Reactive Protein: **Hs-CRP**, Creatine Phosphokinase: **CPK**, Homeostatic Model Assessment for Insulin Resistance index: **HOMA-IR**, Quantitative Insulin Sensitivity Check Index : **QUICKI**, Mid-Arm Circumference: **MAC**, Waist Circumference : **WC ,** Hip Circumference : **HC** .

Acknowledgements

This work was supported by the National Institute for Medical Research Development (NIMAD), Tehran- Iran (no. 957705).

References

1. Tabrizi R, Moosazadeh M, Akbari M, Dabbaghmanesh MH, Mohamadkhani M, Asemi Z, et al. High prevalence of vitamin D deficiency among Iranian population: a systematic review and meta-analysis. Iranian journal of medical sciences. 2018 Mar;43(2):125.

2. Khayyatzadeh SS, Vatanparast H, Avan A, Bagherniya M, Bahrami A, Kiani MA, et al. Serum transaminase concentrations and the presence of irritable bowel syndrome are associated with serum 25-hydroxy vitamin D concentrations in adolescent girls who are overweight and obese. Annals of Nutrition and Metabolism. 2018 Jan 11;71(3-4):234-41.

3. Jean G, Souberbielle JC, Chazot C. Vitamin D in chronic kidney disease and dialysis patients. Nutrients. 2017 Mar 25;9(4):328.

4. Sahota O. Understanding vitamin D deficiency. Age and ageing. 2014 Jul 28;43(5):589-91.

5. Timar A, Saberi-Karimian M, Ghazizadeh H, Parizadeh SM, Sabbaghzadeh R, Emadzadeh M, et al. Evaluation of the serum prooxidant-antioxidant balance before and after vitamin D supplementation in adolescent Iranian girls. Advances in Medical Sciences. 2019 Mar 1;64(1):174-80.

6. Vatandost S, Jahani M, Afshari A, Amiri MR, Heidarimoghadam R, Mohammadi Y. Prevalence of vitamin D deficiency in Iran: a systematic review and meta-analysis. Nutrition and health. 2018 Dec;24(4):269-78.

7. Rodríguez-Rodríguez E, Vizuete AA, Sánchez-Rodríguez P, Mora AM, López-Sobaler AM, Ortega RM. Vitamin D deficiency in Spanish population. Importance of egg on nutritional improvement. Nutricion hospitalaria. 2019 Aug 27;36(Spec No3):3-7.

8. Holick MF, Binkley NC, Bischoff-Ferrari HA, Gordon CM, Hanley DA, Heaney RP, et al. Evaluation, treatment, and prevention of vitamin D deficiency: An endocrine society clinical practice guideline (Journal of Clinical Endocrinology and Metabolism (2011) 96,(1911-1930)). Journal of Clinical Endocrinology and Metabolism. 2011 Dec 1;96(12).

9. Holick MF. The vitamin D deficiency pandemic: Approaches for diagnosis, treatment and prevention. Reviews in Endocrine and Metabolic Disorders. 2017 Jun;18:153-65.

10. Wanner M, Richard A, Martin B, Linseisen J, Rohrmann S. Associations between objective and self-reported physical activity and vitamin D serum levels in the US population. Cancer Causes & Control. 2015 Jun;26:881-91.

11. Ramírez-Prada D, de la Torre MJ, Llórente-Cantarero FJ, Pérez-Navero JL, Gil-Campos M. Evaluation of solar exposure, intake and physical activity in relation with vitamin D serum status in Spanish prepubertal girls. Nutricion Hospitalaria. 2012 Nov 1;27(6):1993-8.

12. Lanske B, Razzaque MS. Vitamin D and aging: old concepts and new insights. The Journal of nutritional biochemistry. 2007 Dec 1;18(12):771-7.

13. Gallagher JC. Vitamin D and aging. Endocrinology and Metabolism Clinics. 2013 Jun 1;42(2):319-32.

14. Jang H, Lee Y, Park K. Obesity and vitamin D insufficiency among adolescent girls and young adult women from Korea. Nutrients. 2019 Dec 13;11(12):3049..

15. Oliai Araghi S, van Dijk SC, Ham AC, Brouwer-Brolsma EM, Enneman AW, Sohl E, et al. BMI and body fat mass is inversely associated with vitamin D levels in older individuals. The journal of nutrition, health & aging. 2015 Dec;19:980-5.

16. Tabesh M, Callegari ET, Gorelik A, Garland SM, Nankervis A, Subasinghe AK, et al, YFHI and Safe-D study groups. Associations between 25-hydroxyvitamin D levels, body composition and metabolic profiles in young women. European Journal of Clinical Nutrition. 2018 Aug;72(8):1093-102..

17. Vitezova A, Muka T, Zillikens MC, Voortman T, Uitterlinden AG, Hofman A, et al. Vitamin D and body composition in the elderly. Clinical Nutrition. 2017 Apr 1;36(2):585-92.

18. De Felice F, Crocetti D, Parisi M, Maiuri V, Moscarelli E, Caiazzo R, et al. Decision tree algorithm in locally advanced rectal cancer: an example of over-interpretation and misuse of a machine learning approach. Journal of cancer research and clinical oncology. 2020 Mar;146:761-5.
19. Connor CW. Artificial intelligence and machine learning in anesthesiology. Anesthesiology. 2019 Dec 1;131(6):1346-59.
20. Esmaeily H, Tayefi M, Ghayour-Mobarhan M, Amirabadizadeh A. Comparing three data mining algorithms for identifying the associated risk factors of type 2 diabetes. Iranian biomedical journal. 2018 Sep;22(5):303.
21. Ghazizadeh H, Shakour N, Ghoflchi S, Mansoori A, Saberi-Karimiam M, Rashidmayvan M, et al. Use of data mining approaches to explore the association between type 2 diabetes mellitus with SARS-CoV-2. BMC Pulmonary Medicine. 2023 Dec;23(1):1-4.
22. Saberi-Karimian M, Mansoori A, Bajgiran MM, Hosseini ZS, Kiyoumarsioskouei A, Rad ES, et al. Data mining approaches for type 2 diabetes mellitus prediction using anthropometric measurements. Journal of Clinical Laboratory Analysis. 2023 Jan;37(1):e24798.
23. Saberi-Karimian M, Safarian-Bana H, Mohammadzadeh E, Kazemi T, Mansoori A, Ghazizadeh H, et al. A pilot study of the effects of crocin on high-density lipoprotein cholesterol uptake capacity in patients with metabolic syndrome: A randomized clinical trial. BioFactors. 2021 Nov;47(6):1032-41.
24. Abdar M. Using decision trees in data mining for predicting factors influencing of heart disease. Carpathian Journal of Electronic and Computer Engineering. 2015 Jul 1;8(2):31.
25. Senanayake S, White N, Graves N, Healy H, Baboolal K, Kularatna S. Machine learning in predicting graft failure following kidney transplantation: A systematic review of published predictive models. International journal of medical informatics. 2019 Oct 1;130:103957.
26. Mansoori A, Hosseini ZS, Ahari RK, Poudineh M, Rad ES, Zo MM, et al. Development of Data Mining Algorithms for Identifying the Best Anthropometric Predictors for Cardiovascular Disease: MASHAD Cohort Study. High Blood Pressure & Cardiovascular Prevention. 2023 May;30(3):243-53.
27. Mansoori A, Sahranavard T, Hosseini ZS, Soflaei SS, Emrani N, Nazar E, et al. Prediction of type 2 diabetes mellitus using hematological factors based on machine learning approaches: a cohort study analysis. Scientific Reports. 2023 Jan 12;13(1):663.
28. Saberi-Karimian M, Mansoori A, Ghazizadeh H, Ferns GA, Aghasizadeh M, Ghayour-Mobarhan M. The Effects of Crocin on HDL Cholesterol Uptake Capacity in Patients with Metabolic Syndrome: A Randomized Clinical Trial. Journal of Nutrition, Fasting & Health. 2022 Jan 2;10.
29. Ogunleye A, Wang QG. XGBoost model for chronic kidney disease diagnosis. IEEE/ACM transactions on computational biology and bioinformatics. 2019 Apr 17;17(6):2131-40.
30. Sharifan P, Bagherniya M, Bajgiran MM, Safarian M, Vatanparast H, Eslami S, et al. The efficacy of dairy products fortified with nano-encapsulated vitamin D3 on physical and mental aspects of the health in obese subjects; the protocol of the SUVINA trial. Translational Metabolic Syndrome Research. 2021 Dec 1;4:1-9.
31. Alberti KG, Zimmet P, Shaw J. Metabolic syndrome—a new world-wide definition. A consensus statement from the international diabetes federation. Diabetic medicine. 2006 May;23(5):469-80.
32. Baecke JA, Burema J, Frijters JE. A short questionnaire for the measurement of habitual physical activity in epidemiological studies. The American journal of clinical nutrition. 1982 Nov 1;36(5):936-42.
33. Etemad Z, Esmailnasab N. The relationship between the level of physical activity and some risk factors of coronary heart disease in the university students. Scientific Journal of Kurdistan University of Medical Sciences. 2012;17(1).
34. Asadi Z, Ghaffarian Zirak R, Yaghooti Khorasani M, Saedi M, Parizadeh SM, Jafarzadeh-Esfehani R, et al. Dietary inflammatory index is associated with healthy eating index, alternative healthy eating index, and dietary patterns among Iranian adults. Journal of Clinical Laboratory Analysis. 2020 Dec;34(12):e23523.
35. Norton K, Olds T, editors. Anthropometrica: a textbook of body measurement for sports and health courses. UNSW press; 1996.
36. Mahler RJ, Adler ML. Type 2 diabetes mellitus: update on diagnosis, pathophysiology, and treatment. The Journal of Clinical Endocrinology & Metabolism. 1999 Apr 1;84(4):1165-71.
37. Sharifan P, Ziaee A, Darroudi S, Rezaie M, Safarian M, Eslami S, et al. Effect of low-fat dairy products fortified with 1500IU nano encapsulated vitamin D3 on cardiometabolic indicators in adults with abdominal obesity: A total blinded randomized controlled trial. Current Medical Research and Opinion. 2021 Apr 3;37(4):579-88.
38. PĂdureanu R, Albu CV, PĂdureanu V, BugĂ AM. Oxidative stress and vitamin D as predictors in multiple sclerosis. Current Health Sciences Journal. 2020 Oct;46(4):371.
39. Ha CD, Cho JK, Lee SH, Kang HS. Serum vitamin D, physical activity, and metabolic risk factors in Korean children. Medicine and science in sports and exercise. 2013 Jan 1;45(1):102-8.

40. Mortensen C, Mølgaard C, Hauger H, Kristensen M, Damsgaard CT. Sun behaviour and physical activity associated with autumn vitamin D status in 4–8-year-old Danish children. Public health nutrition. 2018 Nov;21(17):3158-67.
41. Bauer A, Lechner I, Auer M, Berger T, Bsteh G, Di Pauli F, et al. Influence of physical activity on serum vitamin D levels in people with multiple sclerosis. Plos one. 2020 Jun 11;15(6):e0234333.
42. Von Berens Å, Cederholm T, Fielding RA, Gustafsson T, Kirn D, Laussen J, et al. Physical performance and serum 25 (OH) vitamin D status in community dwelling old mobility limited adults: a cross-sectional study. The journal of nutrition, health & aging. 2018 Jan;22:1-7.
43. Elizondo-Montemayor L, Castillo EC, Rodríguez-López C, Villarreal-Calderón JR, Gómez-Carmona M, Tenorio-Martínez S, et al. Seasonal variation in vitamin D in association with age, inflammatory cytokines, anthropometric parameters, and lifestyle factors in older adults. Mediators of inflammation. 2017 Sep 14;2017.
44. Meehan M, Penckofer S. The role of vitamin D in the aging adult. Journal of aging and gerontology. 2014 Dec;2(2):60.
45. Wimalawansa SJ. Vitamin D deficiency: effects on oxidative stress, epigenetics, gene regulation, and aging. Biology. 2019 May 11;8(2):30.
46. Leiu KH, Chin YS, Mohd Shariff Z, Arumugam M, Chan YM. High body fat percentage and low consumption of dairy products were associated with vitamin D inadequacy among older women in Malaysia. PLoS One. 2020 Feb 13;15(2):e0228803.
47. Fatima SS, Farooq S, Tauni MA, Irfan O, Alam F. Effect of raised body fat on vitamin D, leptin and bone mass. Journal of Pakistan Medical Association. 2015;65(12):1315.
48. Ong MW, Tan CH, Cheng AK. Prevalence and determinants of vitamin D deficiency among the overweight and obese Singaporeans seeking weight management including bariatric surgery: a relationship with bone health. Obesity Surgery. 2018 Aug;28:2305-12.
49. Shantavasinkul PC, Phanachet P, Puchaiwattananon O, Chailurkit LO, Lepananon T, Chanprasertyotin S, et al. Vitamin D status is a determinant of skeletal muscle mass in obesity according to body fat percentage. Nutrition. 2015 Jun 1;31(6):801-6.
50. Grønborg IM, Lundby IM, Mølgaard C, Jakobsen J, Ovesen L, Andersen R. Association of body fat and vitamin D status and the effect of body fat on the response to vitamin D supplementation in Pakistani immigrants in Denmark. European Journal of Clinical Nutrition. 2015 Mar;69(3):405-7.
51. Wali S, Alsafadi S, Abaalkhail B, Ramadan I, Abulhamail B, Kousa M, et al. The association between vitamin D level and restless legs syndrome: a population-based case-control study. Journal of Clinical Sleep Medicine. 2018 Apr 15;14(4):557-64.
52. Balaban H, Yıldız ÖK, Çil G, Şentürk İA, Erselcan T, Bolayır E, et al. Serum 25-hydroxyvitamin D levels in restless legs syndrome patients. Sleep medicine. 2012 Aug 1;13(7):953-7.
53. Wali S, Shukr A, Boudal A, Alsaiari A, Krayem A. The effect of vitamin D supplements on the severity of restless legs syndrome. Sleep and breathing. 2015 May;19:579-83.
54. Romano F, Muscogiuri G, Di Benedetto E, Zhukouskaya VV, Barrea L, Savastano S, et al. Vitamin D and sleep regulation: is there a role for vitamin D?. Current pharmaceutical design. 2020 Jun 1;26(21):2492-6.
55. Wali SO, Abaalkhail B, Alhejaili F, Pandi-Perumal SR. Efficacy of vitamin D replacement therapy in restless legs syndrome: a randomized control trial. Sleep and Breathing. 2019 Jun 1;23:595-601.