JOURNAL OF CLINICAL MEDICINE OF KAZAKHSTAN (E-ISSN 2313-1519)
Review Article DOI: https://doi.org/10.23950/jcmk/18013

Determination of Body Mass Index Using Machine Learning Regression Methods from Body Composition Parameters

Seda Sertel Meyvaci[1], Yusuf Secgin[2], Taha Gokmen Ulger[3], Tuba Taslamacioglu Duman[4], Beyza Celik[1], Sena Demiroglu[1]

[1]*Department of Anatomy, Faculty of Medicine, Bolu Abant Izzet Baysal University, Bolu, Turkey*
[2]*Department of Anatomy, Faculty of Medicine, Karabük University, Karabük, Turkey*
[3]*Department of Nutrition and Dietetics, Faculty of Health Sciences, Bolu Abant Izzet Baysal University, Bolu, Turkey*
[4]*Department of Internal Medicine, Faculty of Medicine, Bolu Abant Izzet Baysal University, Bolu, Turkey*

Received: 2025-10-14.
Accepted: 2026-02-21.

J Clin Med Kaz 2026; 23(2): 45-50

Corresponding author:
Sena Demıroglu
E-mail: demiroglusenaa@gmail.com.
ORCID: 0000-0002-7214-5538.

ABSTRACT

Introduction: Obesity is one of the most significant global health problems increasing today. Body mass index (BMI) is a fundamental parameter widely used to assess obesity. However, evaluating different measurements of body composition may contribute to determining BMI more accurately and comprehensively. In this context, machine learning (ML) provides a powerful alternative to classical methods by predicting BMI from body composition. Based on this hypothesis, the aim of this study is to predict BMI using ML regression models from body composition parameters.

Methods: The study included 411 obese individuals aged 18–65 years. The individuals' body weight and body composition parameters [fat mass index, body fluid percentage, fat mass (kg), fat-free mass (kg), body fat percentage and fat-free mass percentage] were measured using the Tanita MC 580 body composition analyzer. Data were analyzed using 16 different ML regression models with the Python language and PyCaret library. Model performance was primarily evaluated using the coefficient of determination (R^2).

Results: The highest success was achieved with Linear Regression, Bayesian Ridge, and Ridge Regression models (R^2=0.9937). Huber Regressor (0.9928) and Least Angle Regression (0.9911) followed. Among the models with low success was K Neighbors Regressor (0.8316). In the parameter analysis, Fat-Mass Index was the strongest predictor in BMI estimation.

Conclusion: This study shows that ML-based regression models can predict BMI with high accuracy using body composition parameters, with the fat mass index providing the most significant contribution. These results emphasize the potential of digital phenotyping as a valuable approach in obesity research and clinical evaluation. By framing BMI not only as a traditional measurement but as a more comprehensive health marker supported by digital tools, the findings point toward the development of personalized, rapid, and reliable digital health solutions for obesity screening and monitoring.

Keywords: Body mass index, obesity, machine learning regression models, bioelectrical impedance analysis, digital phenotyping.

Introduction

Obesity is a complex public health problem with rapidly increasing prevalence and is associated with numerous health conditions such as cardiovascular diseases, type 2 diabetes, osteoarthritis, and certain types of cancer [1]. Among the most commonly used indicators for the classification of obesity, body mass index (BMI) is widely utilized in both clinical and epidemiological studies due to its ease of application and simple calculation method. However, since BMI is solely based on height and body weight measurements, it cannot reflect body fat distribution, the distinction between fat mass and fat-free mass, or metabolically risky abdominal obesity [2-4]. This limitation may lead to misclassification, particularly in individuals with high muscle mass, the elderly, or those with different body composition profiles [5].

To overcome these limitations, bioelectrical impedance analysis (BIA) provides an important alternative. BIA offers detailed information on an individual's body composition by safely, non-invasively, and reproducibly determining parameters such as fat mass, fat-free mass, and body fluid [6]. In clinical practice, BIA-derived data have been shown to provide more reliable results than BMI in predicting cardiometabolic risks, accurately characterizing obesity phenotypes, and developing personalized nutritional approaches [7-9].

Meanwhile, machine learning (ML) methods, which have gained increasing importance in health sciences in recent years, enable the modeling of complex relationships among multidimensional biological data. ML algorithms can process large datasets to uncover patterns that may be overlooked by traditional methods, thereby contributing to a more precise and individualized assessment of obesity and related metabolic disorders [10, 11]. Integrating BIA-derived body composition parameters with ML techniques not only provides more accurate predictions beyond BMI but also offers practical advantages for clinical applications [12]. Machine learning has also been applied to predict temporomandibular disorders using clinical parameters and to predict carpal tunnel syndrome using anthropometric and strength-based measurements [13, 14]. Although machine learning has been widely applied for obesity classification, few studies have focused exclusively on BMI prediction using BIA-derived body composition parameters.

To the best of our knowledge, no previous study has directly compared 16 machine learning regression algorithms for BMI prediction using exclusively BIA-derived body composition parameters. In this context, the present study systematically evaluates and compares a wide range of machine learning regression models to estimate BMI using biologically meaningful body composition parameters obtained from BIA. Thus, it is intended to develop a model that overcomes the limitations of conventional BMI calculation, incorporates body composition, and allows for more personalized assessments in clinical practice. By emphasizing both predictive performance and model interpretability, this study aims to clarify the relationship between BMI and underlying body composition and to highlight the potential clinical relevance of body composition–based obesity assessment.

Methods

This study included patients with obesity, aged 18–65 years, with a BMI of ≥30 kg/m², who presented to the Internal Medicine Outpatient Clinic of Bolu Abant Izzet Baysal Training and Research Hospital and were referred to the diet outpatient clinic. Body composition data of these patients were obtained using a Tanita MC 580 BIA device. The measurements, together with demographic information recorded for clinical follow-up, were retrospectively retrieved from outpatient clinic records. For patients with multiple measurements, only the first measurement was considered. The body composition parameters evaluated were fat mass, body fat percentage, fat-free mass, fat-free mass percentage, fat mass index (FMI), total body water (% of body weight) and BMI. Both FMI and BMI values were automatically calculated by the Tanita MC 580 BIA device using fat mass, height, and body weight inputs. FMI was derived by dividing fat mass in kilograms by the square of height in meters [fat mass (kg) / height² (m²)], whereas BMI was calculated by dividing body weight in kilograms by the square of height in meters [body weight (kg) / height² (m²)]. Ethical approval for the study was obtained from the Bolu Abant Izzet Baysal University Non-Interventional Clinical Research Ethics Committee (2025/384).

Machine Learning Regression Models

To estimate BMI, the Python programming language and the PyCaret library were employed. Modeling procedures were carried out on a Monster Abra A7 computer with an Intel i5 processor. The collected numerical data (body composition parameters and age) constituted the input layer for ML algorithms and regression models, while the output layer was designed as the predicted BMI.

A total of 16 PyCaret ML regression models were included in the study: Linear Regression (LiR), Bayesian Ridge Regression (BRR), Ridge Regression (RR), Huber Regression (HR), Least Angle Regression (LAR), Extra Tree Regression (ETR), Elastic Net (EN), Gradient Boosting Regression, Lasso Regression (LaR), Lasso Least Angle Regression, Passive Aggressive Regression, Random Forest Regression (RFR), Light Gradient Boosting Machine Regression, Decision Tree Regression (DTR), AdaBoost Regression, and K-Neighbors Regression. LiR is a fundamental method that models the linear relationship between the dependent and independent variables [15]. BRR models the joint probability distribution of variables, thereby allowing the evaluation of event probabilities and related risks [16]. Among regularization methods, RR reduces multicollinearity by adding penalties to parameter magnitudes [17], while LaR shrinks regression coefficients and reduces some to zero, enabling variable selection [18]. EN combines the advantages of Lasso and Ridge to perform both variable selection and multicollinearity reduction [19]. Additionally, HR is preferred as a method that is less affected by outliers [20]. LARS is used as an algorithm for variable selection in high-dimensional datasets [21].

For model development, the dataset was divided into 80% training and 20% test sets. The training set was used to train the models, while the test set was used for independent performance evaluation. To mitigate potential overfitting concerns associated with high R^2 values, model performance metrics were assessed separately on the training and test datasets. Hyperparameter tuning was performed to optimize model performance. The evaluated performance metrics included Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Root Mean Squared Logarithmic Error (RMSLE), Mean Absolute Percentage Error (MAPE), Total Taken (TT-seconds), and the Coefficient of Determination (R^2). Total Taken (TT-seconds) refers to the total elapsed time required to train and evaluate each regression model.

To improve model interpretability, SHAP (SHapley Additive exPlanations) analysis was applied to identify the relative contribution of each input parameter to BMI prediction in the best-performing models.

Basic Statistical Analysis

The normality of the parameters was assessed using the Anderson–Darling test, and descriptive statistics of the data were reported. All analyses were performed using the Minitab 17 statistical software package.

Results

A total of 411 participants aged 18–65 years were included in the study. The mean age of the participants was 38 years (range: 18–65). Descriptive statistics of the body composition parameters and BMI values of the participants are presented in Table 1.

Table 1 Descriptive Statistics of the Parameters

Parameters	Median (Min–Max)
Fat Mass Index (FMI)	15.165 (5.988-35.371)
Total Body Water (%)	39.900 (30.200-71.800)
Fat Mass (kg)	40.000 (16.900-86.900)
Fat-Free Mass (kg)	54.500 (41.200-98.100)
Body Fat Percentage (%)	42.303 (19.537-57.143)
Fat-Free Mass Percentage (%)	57.697 (42.067-80.463)
Body Mass Index (BMI)	36.208 (30.010-65.616)

The performance metrics obtained from the analyses conducted with PyCaret, including MAE, MSE, RMSE, R^2, RMSLE, MAPE, and TT-seconds, are presented in Table 2. Examination of Table 2 revealed that the highest R^2 values were achieved with the LiR, BRR, and RR models. These models yielded an R^2 of 0.9937, demonstrating very high accuracy in BMI prediction. They were followed by the Huber Regressor (0.9928) and LAR (0.9911). Among the models with lower performance, the K-Neighbors Regressor (0.8316) was noted.

Figure 1 presents the prediction error and residual plots obtained after hyperparameter optimization (tuning) for the best-performing models. Examination of Figure 1 shows that the plots indicate the models explained the majority of the variance and that the predictions exhibited a high degree of agreement with the actual values.

Figure 2 illustrates the importance of body composition parameters contributing most to the prediction performance of the models with the highest R^2 values. Examination of Figure 2 revealed that FMI was the most influential variable in BMI prediction.

Table 2 Performance Metrics of Machine Learning Regression Models

Models	MAE	MSE	RMSE	R2	RMSLE	MAPE	TT-seconds
Linear Regression	0.3630	0.2582	0.4877	0.9937	0.0121	0.0095	0.3330
Bayesian Ridge Regression	0.3633	0.2596	0.4882	0.9937	0.0122	0.0095	0.0100
Ridge Regression	0.3634	0.2604	0.4881	0.9937	0.0122	0.0095	0.0110
Huber Regressor	0.3575	0.2950	0.5193	0.9928	0.0127	0.0092	0.0120
Least Angle Regression	0.3855	0.3539	0.5580	0.9911	0.0137	0.0100	0.0110
Extra Trees Regressor	0.7869	1.4198	1.1697	0.9633	0.0279	0.0201	0.0440
Elastic Net	0.9675	1.4658	1.2000	0.9616	0.0310	0.0258	0.0100
Gradient Boosting Regressor	0.9100	1.5680	1.2143	0.9602	0.0288	0.0233	0.0270
Lasso Regression	1.0381	1.6867	1.2883	0.9557	0.0334	0.0278	0.2950
Lasso Least Angle Regression	1.0386	1.6884	1.2889	0.9557	0.0334	0.0278	0.0120
Passive Aggressive Regressor	1.0597	1.7959	1.3167	0.9512	0.0335	0.0280	0.0110
Random Forest Regressor	1.0656	2.2939	1.4728	0.9414	0.0346	0.0268	0.0530
Light Gradient Boosting Machine	1.2405	3.5113	1.8254	0.9118	0.0400	0.0303	0.0500
Decision Tree Regressor	1.3203	3.5875	1.8423	0.9116	0.0431	0.0334	0.0110
AdaBoost Regressor	1.5018	3.6011	1.8650	0.9068	0.0463	0.0395	0.0220
K-Neighbors Regressor	2.0024	6.2883	2.4632	0.8316	0.0607	0.0521	0.0120

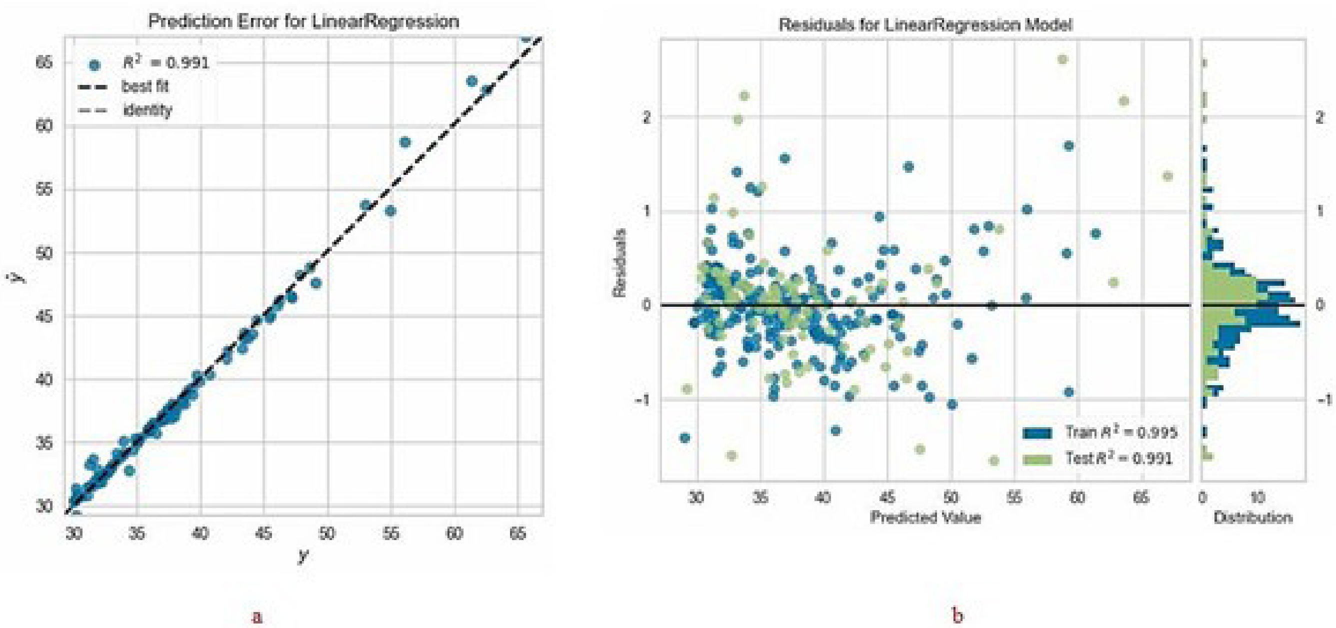

Figure 1 – Plots of the Best-Performing Models (a: Prediction Error, b: Residuals Plot)

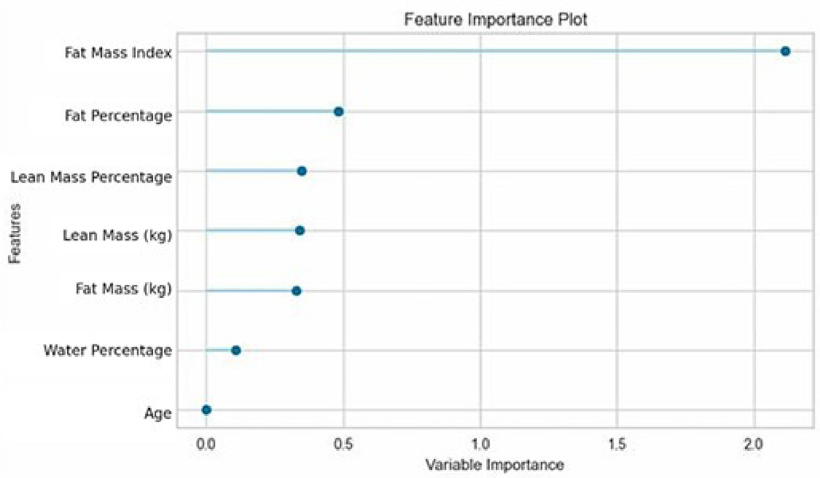

Figure 2 – Importance Plot of Key Parameters

Discussion

In this study, the point prediction of BMI was examined using ML-based regression models derived from body composition parameters obtained through the BIA method. The analyses demonstrated that the LiR, BRR, and RR models achieved the highest accuracy rates, indicating that BMI could be reliably predicted within the given mathematical framework. Notably, the strong performance of linear-based regression models suggests that the relationship between BMI and body composition parameters is predominantly linear and biologically meaningful. This finding provides conceptual insight, indicating that complex non-linear machine learning architectures may not be necessary when clinically relevant body composition predictors are used. These findings suggest that ML methods may provide complementary and supportive insights alongside traditional approaches in the assessment and monitoring of obesity. It should be emphasized that the purpose of the regression analysis applied in this study is not to replicate the BMI formula, but to evaluate the contribution of the parameters used, including those directly related to fat mass index (FMI) and fat-free mass index (FFMI) derived from body weight. Therefore, SHAP analysis was incorporated to enhance model interpretability by demonstrating the relative contribution of each parameter to the model predictions.

The results further revealed that FMI was the most influential parameter in BMI prediction. This result supports the view that BMI fundamentally reflects fat mass–related characteristics rather than overall body weight alone, reinforcing the clinical importance of integrating fat-based indices into obesity assessment. This finding highlights the central role of fat mass in the evaluation of obesity and underscores its importance in personalized healthcare approaches. Indeed, the literature emphasizes the strong association between body fat mass, body fat percentage and metabolic health, particularly with respect to insulin resistance, lipid profiles, and cardiometabolic risk factors [1, 22]. Moreover, FMI has been shown to be clinically valuable not only in obesity classification but also in identifying “metabolically obese but normal-weight” individuals [22]. These results demonstrate that body composition parameters may provide more accurate classifications in clinical decision-making processes than conventional BMI measurements.

Recent studies in the literature also support the potential of ML approaches to contribute multidimensionally to obesity assessment. Delnevo et al. examined BMI prediction using psychological indicators and demonstrated that the presence of depression significantly influenced the predictions, emphasizing the importance of considering psychosocial variables in obesity evaluation [23]. In a study by La Cruz et al., 24 different anthropometric measures were used to classify individuals according to body fat percentage through support vector machines, yielding high accuracy and sensitivity values; however, the limited specificity highlighted the likelihood of false positives [24]. Similarly, Genç & Arıcan incorporated demographic variables in addition to conventional BMI components for BMI classification and reported that the random forest algorithm achieved the highest performance [25]. Chen et al., integrated clinical and lifestyle data in childhood obesity and demonstrated that ML methods could capture the multifactorial nature of the condition, providing applicable tools for clinical decision-support systems [26]. Collectively, these studies indicate that ML models can operate effectively with diverse data types and hold broad potential in obesity evaluation.

Evidence also suggests that body composition data can be utilized to predict the success of treatment protocols [27, 28]. Nevertheless, the literature also shows that the relationship between ML-derived clinical predictions and BMI remains highly complex. Therefore, future studies should aim to integrate psychological, demographic, and lifestyle variables alongside body composition parameters to model BMI in a more accurate and clinically meaningful way. The originality of this study lies in the systematic comparison of different ML regression methods for point prediction of BMI using exclusively body composition parameters obtained through BIA. Rather than representing a purely technical machine learning application, this approach contributes to a clearer scientific understanding of how BMI relates to its underlying body composition determinants, thereby strengthening the conceptual and clinical relevance of BMI-based evaluations. By demonstrating the predictive performance of these regression models, the study provides concrete evidence of their applicability in clinical prediction and proposes a practical approach to overcoming the limitations of conventional BMI calculations. Furthermore, the potential integration of these ML models into routine clinical practice should be considered, highlighting how they may support personalized decision-making in obesity management. In this respect, the research offers methodological innovation to the literature and provides practical contributions to clinical practice.

Limitations and Future Directions

This study has several limitations. Despite the high accuracy observed, the potential risk of overfitting and the limited generalizability of our findings to other populations should be considered. First, the analyses were restricted to body composition parameters obtained solely through BIA; incorporating additional anthropometric, biochemical, and psychosocial variables could potentially improve prediction accuracy. Moreover, the sample size and demographic characteristics may limit the generalizability of the findings. Study limitations include the fact that the research was conducted in a single population at a single center, and non-BMI variables such as biochemical, demographic, and lifestyle factors were not included. Therefore, future research should consider conducting similar analyses with larger and more diverse samples, encompassing different age groups, sexes, and ethnic backgrounds.

In addition, adopting longitudinal study designs would allow the monitoring of BMI changes over time and provide an opportunity to evaluate the long-term predictive power of the models. Finally, future studies should not only focus on the statistical accuracy of ML models but also address their clinical applicability, interpretability, and usability by healthcare professionals, which should be among the primary objectives of forthcoming research.

Conclusion

In this study, point prediction of BMI was performed using ML-based regression models with body composition parameters obtained through BIA (fat mass, body fat percentage, fat-free mass, fat-free mass percentage, FMI, total body water (% of body water), and BMI). The LiR, BRR, and RR algorithms achieved the highest R^2 values, providing accurate BMI predictions and explaining a large proportion of variance. Moreover, FMI was identified as the most influential predictor. The strong performance of linear-based regression models suggests that the relationship between BMI and its underlying body composition determinants is largely linear and biologically interpretable. This finding provides conceptual insight rather than a purely technical outcome, emphasizing the scientific relevance of body composition–based BMI modeling.

These findings indicate that, independent of conventional BMI calculation, ML-based models can show promising predictive performance from body composition data, offering a practical and efficient tool for obesity assessment, and may contribute to the future development of AI-assisted personalized healthcare applications. In this respect, the study highlights the importance of interpreting BMI in conjunction with body composition parameters rather than as an isolated anthropometric index.

Future studies should address the limitations identified in the present work by incorporating non-BMI variables, such as biochemical, demographic, and lifestyle factors, and by conducting analyses across different centers and diverse populations. Such efforts will enhance the clinical applicability, generalizability, and robustness of ML-based BMI prediction models. By focusing on both predictive performance and interpretability, the present study contributes to a more clinically meaningful and scientifically grounded use of machine learning in obesity research.

Author Contributions: Conceptualization, S. S. M., Y. S.; methodology / planning and organization, Y. S., S. S. M.; funding acquisition, S. S. M.; writing – original draft preparation, S. S. M., T. G. U., T. T. D., B. C., S. D.; materials, S. S. M., Y. S., T. G. U., T. T. D.; data collection, S. S. M., Y. S., T. G. U., T. T. D., B. C., S. D.; data analysis and statistics, Y. S. All authors have read and approved the final version of the manuscript.

Disclosures: The authors have no conflicts of interest.

Acknowledgments: None.

Funding: None.

Data availability statement: The corresponding author can provide the data supporting the study's conclusions upon request. Due to ethical and privacy constraints, the data are not publicly accessible.

Artificial Intelligence (AI) Disclosure Statement: The authors declare no AI Tools used for preparation of this work.

References

1. Blüher M. Metabolically healthy obesity. Endocr Rev. 2020;41(3):bnaa004. https://doi.org/10.1210/endrev/bnaa004. PMID: 32128581; PMCID: PMC7098708.
2. Jeong SM, Lee DH, Rezende LFM, Giovannucci EL. Different correlation of body mass index with body fatness and obesity-related biomarker according to age, sex and race-ethnicity. Sci Rep. 2023;13(1):3472. https://doi.org/10.1038/s41598-023-30527-w. PMID: 36859451; PMCID: PMC9977890.
3. Pray R, Riskin S. The history and faults of the body mass index and where to look next: a literature review. Cureus. 2023 Nov 3;15(11):e48230. https://doi.org/10.7759/cureus.48230. PMID: 38050494; PMCID: PMC10693914.
4. Gómez-Ambrosi J, Silva C, Galofré JC, Escalada J, Santos S, Alcaraz-Millán D, Vila N, Ibáñez Ibáñez P, Cancho Gil MJ, Valentí V, Rotellar F, Ramírez B, Salvador J, Frühbeck G. Body mass index classification misses subjects with increased cardiometabolic risk factors related to elevated adiposity. Int J Obes (Lond). 2011;36(2):286–294. https://doi.org/10.1038/ijo.2011.100.
5. Wu Y, Li D, Vermund SH. Advantages and limitations of the body mass index (BMI) to assess adult obesity. Int J Environ Res Public Health. 2024;21(6):757. https://doi.org/10.3390/ijerph21060757.
6. Son JW, Han BD, Bennett JP, Heymsfield S, Lim S. Development and clinical application of bioelectrical impedance analysis method for body composition assessment. Obes Rev. 2025;26(1):e13844. https://doi.org/10.1111/obr.13844.
7. Guo Y, Zhang M, Ye T, Wang Z, Yao Y. Application of bioelectrical impedance analysis in nutritional management of patients with chronic kidney disease. Nutrients. 2023;15(18):3941. https://doi.org/10.3390/nu15183941.

8. Böhm A, Heitmann BL. The use of bioelectrical impedance analysis for body composition in epidemiological studies. Eur J Clin Nutr. 2013;67(Suppl 1):S79–85. https://doi.org/10.1038/ejcn.2012.168. PMID: 23299875.
9. Kobayashi J, Murano S, Kawamura I, Nakamura F, Murase Y, Kawashiri MA, Nohara A, Asano A, Inazu A, Mabuchi H. The relationship of percent body fat by bioelectrical impedance analysis with blood pressure, and glucose and lipid parameters. J Atheroscler Thromb. 2006;13(5):221–6. https://doi.org/10.5551/jat.13.221. PMID: 17146149.
10. Nath T, Ahima RS, Santhanam P. Body fat predicts exercise capacity in persons with type 2 diabetes mellitus: a machine learning approach. PLoS One. 2021;16(3):e0248039. https://doi.org/10.1371/journal.pone.0248039.
11. DeGregory KW, Kuiper P, DeSilvio T, Pleuss JD, Miller R, Roginski JW, Fisher CB, Harness D, Viswanath S, Heymsfield SB, Dungan I, Thomas DM. A review of machine learning in obesity. Obes Rev. 2018;19(5):668–685. https://doi.org/10.1111/obr.12667. PMID: 29426065; PMCID: PMC8176949.
12. Mohammed A. Predicting heart disease with body composition using a hybrid machine learning approach. Res Sq. 2025;July:rs-7041567/v1. https://doi.org/10.21203/rs.3.rs-7041567/v1.
13. Yıldız NT, Kocaman H, Yıldırım H, Canlı M. An investigation of machine learning algorithms for prediction of temporomandibular disorders by using clinical parameters. Medicine. 2024;103(41):e39912. https://doi.org/10.1097/MD.0000000000039912.
14. Yetiş M, Kocaman H, Canlı M, Yıldırım H, Yetiş A, Ceylan İ. Carpal tunnel syndrome prediction with machine learning algorithms using anthropometric and strength-based measurement. PLoS One. 2024;19(4):e0300044. https://doi.org/10.1371/journal.pone.0300044.
15. James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning: with applications in R. 2nd ed. New York: Springer; 2021. 607 s.
16. Bottigliengo D, Berchialla P, Lanera C, Azzolina D, Lorenzoni G, Martinato M, Giachino D, Baldi I, Gregori D. The role of genetic factors in characterizing extra-intestinal manifestations in Crohn's disease patients: are Bayesian machine learning methods improving outcome predictions? J Clin Med. 2019;8(6):865. https://doi.org/10.3390/jcm8060865. PMID: 31212952; PMCID: PMC6617350.
17. Hoerl AE, Kennard RW. Ridge regression: biased estimation for nonorthogonal problems. Technometrics. 1970;12:55–67. https://doi.org/10.1080/00401706.1970.10488634.
18. Tibshirani R. Regression shrinkage and selection via the Lasso. J R Stat Soc Ser B Methodol. 1996;58(1):267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x.
19. Zou H, Hastie T. Regularization and variable selection via the elastic net. J R Stat Soc Ser B Stat Methodol. 2005;67(2):301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x.
20. Huber PJ. Robust estimation of a location parameter. Ann Math Stat. 1964;35:73–101. https://doi.org/10.1214/aoms/1177703732.
21. Efron B, Hastie T, Johnstone I, Tibshirani R. Least angle regression. Ann Stat. 2004;32(2):407–499. https://doi.org/10.1214/009053604000000067.
22. Kim B, Taniguchi K, Isobe T, Oh S. Triglyceride-glucose index is capable of identifying metabolically obese, normal-weight older individuals. J Physiol Anthropol. 2024;22:22. https://doi.org/10.1186/s40101-024-00355-6.
23. Delnevo G, Mancini G, Roccetti M, Salomoni P, Trombini E, Andrei F. The prediction of body mass index from negative affectivity through machine learning: a confirmatory study. Sensors. 2021;21(7):2361. https://doi.org/10.3390/s21072361.
24. Lacruz ALA, Severeyn E, Huerta M, Wong S. Support vector machine technique as classifier of impaired body fat percentage. In: Tallón-Ballesteros AJ, editor. Fuzzy Systems and Data Mining VII. 2021:197–210. https://doi.org/10.3233/FAIA210188.
25. Genc AC, Arıcan E. Obesity classification: a comparative study of machine learning models excluding weight and height data. Rev Assoc Med Bras (1992). 2025;71(1):e20241282. https://doi.org/10.1590/1806-9282.20241282. PMID: 40105561; PMCID: PMC11918863.
26. Chen F, Melton PE, Vinsen K, Mori T, Beilin L, Huang R-C. Explainable AI to predict a complex multifactorial outcome, childhood obesity: application to clinical epidemiology. medRxiv. 2025. https://doi.org/10.1101/2025.06.21.25330041.
27. Khazanchi R, Bajaj A, Shah RM, Chen AR, Reyes SG, Kurapaty SS, Hsu WK, Patel AA, Divi SN. Using machine learning and deep learning algorithms to predict postoperative outcomes following anterior cervical discectomy and fusion. Clin Spine Surg. 2023;36(3):143–149. https://doi.org/10.1097/BSD.0000000000001443. PMID: 36920355.
28. Tahmassebi A, Wengert GJ, Helbich TH, Bago-Horvath Z, Alaei S, Bartsch R, Dubsky P, Baltzer P, Clauser P, Kapetas P, Morris EA, Meyer-Baese A, Pinker K. Impact of machine learning with multiparametric magnetic resonance imaging of the breast for early prediction of response to neoadjuvant chemotherapy and survival outcomes in breast cancer patients. Invest Radiol. 2019;54(2):110–117. https://doi.org/10.1097/RLI.0000000000000518. PMID: 30358693; PMCID: PMC6310100.