Annals of Medicine and Medical Sciences (AMMS)
Volume 04, 2025, *Page No.: 1190 to 1200*
Available at: *http://ammspub.com*

Review Article

Advancements in AI for Obesity Prediction: A Systematic Review and Meta-Analysis

Veerakesari S [1], Sarathlal S [1], Aleena Thomas Cheeran [1], Noula Rahim [2], Jamila Hameed *[3]

[1]Department of General Medicine, Karuna Medical College, Vilayodi, Chittur, Palakkad, Kerala, India.
[2]Department of Biochemistry, Karuna Medical College, Vilayodi, Chittur, Palakkad, Kerala, India.
[3]Department of Obstetrics and Gynaecology, Karuna Medical College, Vilayodi, Chittur, Palakkad, Kerala, India.

*Corresponding Author: Jamila Hameed; *hameedjamila78@gmail.com*

Abstract

Background: The worldwide prevalence of obesity has reached an epidemic level and has presented a significant burden to public health infrastructure. There is a critical need for novel predictive instruments, and computerized intelligent systems (artificial intelligence or AI), specifically machine learning (ML), are an appealing method to precisely forecast obesity and associated health outcomes. Nonetheless, an extensive summary of recent evidence on the efficacy and usage of these models remains a critical void within literature. ***Aim and Objective:*** The principal research question to be addressed is: “What is the overall predictive efficacy of computer and machine-based intelligence systems in the prediction of obesity and overweight status in adult and adolescent populations?” ***Methods:*** Systematic review and meta-analysis were conducted on articles released in the time period 2020 to 2025. The databases used in this review were PubMed, Scopus, Web of Science, and Embase. The relevant literature was searched by using the application of keywords and Boolean operators, namely "obesity," "overweight," "artificial intelligence," "machine learning," "deep learning," and "prediction." The total of ten studies were covered in the systematic review while four studies were chosen to be used in the meta-analysis. The quality of the studies were assessed by using Newcastle–Ottawa Scale (NOS). A meta-analysis was undertaken using a random-effects model to calculate the pooled effect size, standard error, and 95% confidence interval with statistical analysis being done in RStudio. ***Results:*** A total of 307 studies were identified by the database searching. 92 duplicates and 215 abstracts were removed by screening. The full systematic review covered 10 studies. The meta-analysis in four of these studies with a pooled sample size of 363,731 produced a pooled effect size (proportion) of 0.730 (95% CI = 0.719 to 0.741) demonstrating a strong degree of predictive accuracy. ***Conclusion:*** Artificial intelligence and machine learning models consistently exhibit superior predictive capabilities in estimating obesity and overweight conditions. The results underscore the clinical significance of these models as essential instruments for early prevention and intervention efforts, providing an accurate methodology to tackle the public health challenge posed by obesity.

Keywords: ***Artificial intelligence, Machine learning, Obesity, Prediction, Systematic review, Meta-analysis.***

Introduction

The global incidence of obesity has escalated to epidemic levels, resulting in a significant strain on public health infrastructures and economic systems. As reported by the World Health Organization (WHO), the global prevalence of obesity has nearly tripled since 1975, affecting more than 650 million adults. Given its role as a principal risk factor for noncommunicable diseases, including type 2 diabetes, cardiovascular conditions, and certain types of cancer, effective prevention and management strategies are critically necessary. Conventional approaches for evaluating obesity risk, such as the Body Mass Index (BMI) and clinical questionnaires, frequently fail to adequately capture the intricate and multifactorial characteristics of obesity, which stem from a dynamic interaction of genetic, environmental, lifestyle, and behavioral elements. Although these methods are readily available, they may lack the detailed and precise risk assessments essential for delivering efficient preventive care. Therefore, there is a growing interest in utilizing artificial intelligence and machine learning to create more advanced predictive models for obesity (Azmi *et al*., 2025) (Gasmi, 2022). These sophisticated computational techniques have the potential to uncover complex patterns within extensive datasets, thereby improving the accuracy and individualization of obesity risk evaluations (Bhatia *et al*., 2022). The utilization of AI and machine learning transcends mere risk assessment, providing detailed insights into personalized dietary and lifestyle advice, which is particularly vital in light of the increasing prevalence of sedentary habits and unhealthy eating patterns in contemporary lifestyles (Tsolakidis *et al*., 2024). Indeed, the integration of artificial intelligence technologies is becoming essential in the domain of nutrition research, particularly as this field grows to investigate the intricate relationships among food, personal health, and community well-being (Kassem *et al*., 2025).

Over the past few years, the exponential expansion of health information combined with remarkable increases in computational potency has brought forth artificial intelligence (AI) as a central tool in health study and clinical application. Machine learning (ML), an offshoot of AI, offers a sophisticated framework for the exploration

Received: September 03, 2025; **Revised:** October 05, 2025; **Accepted:** October 11, 2025

of massive and highly complex data sets to elicit intricate relations and patterns that often escape classical statistical approaches. Utilization of AI is broad-based across several dimensions of the study of obesity and includes elements ranging from risk identification of disease to individualized tailoring of nutritional and weight-reduction plans. Accumulating literature supports the conclusion that AI types often exhibit greater descriptive accuracy than traditional statistical methods; however, differences in performance exist depending on the specific task, information source, and population subgroup. The aim of this systematic survey and meta-analysis is to gather the present evidence relating to the application of AI and machine learning approaches to obesity prediction and to emphasize key approaches to performance evaluation and emerging trends in this rapidly expanding field.

Although an exponential amount of literature on the usage of AI in obesity has emerged, it is necessary to achieve an exhaustive synthesis of evidence to offer an absolute quantitative overview of these models' effectiveness. As some reviews have inspected individual applications, a high-powered meta-analysis specifically on the predictive ability of AI models in regards to obesity and overweight status has been absent. Focusing on synthesizing this evidence is vital to both clinicians attempting to incorporate these instruments into their toolkit and researchers trying to establish areas of exploration.

This systematic review and meta-analysis seek to bridge this gap by bringing together a combined evaluation of the predictive accuracy of AI and ML models of obesity. Through systematic identification, appraisal, and aggregation of findings of recent evidence, we aim to answer a specific research query: What is the aggregated predictive accuracy of machine learning and artificial intelligence models in predicting status of overweight and obesity in adolescents and adult populations? The results will give a strong evidence base to guide the creation of more accurate and efficient AI-based prediction models and direct their clinical application to reduce the worldwide burden of obesity.

Methodology

Search Strategy

A systematic search was performed across four major electronic databases: PubMed, Scopus, Embase, and Web of Science. The search was limited to publications from January 1, 2020, to September 15, 2025. The following keywords and Boolean operators were used to construct the search strings: ("obesity" OR "overweight") AND ("artificial intelligence" OR "machine learning" OR "deep learning") AND ("prediction" OR "predictive model" OR "risk score").

Study Design

This study followed a systematic review and meta-analysis design, adhering to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines (Figure 1). A systematic review was conducted on all relevant studies, while a meta-analysis was performed on a subset of studies that provided quantitative data suitable for pooling.

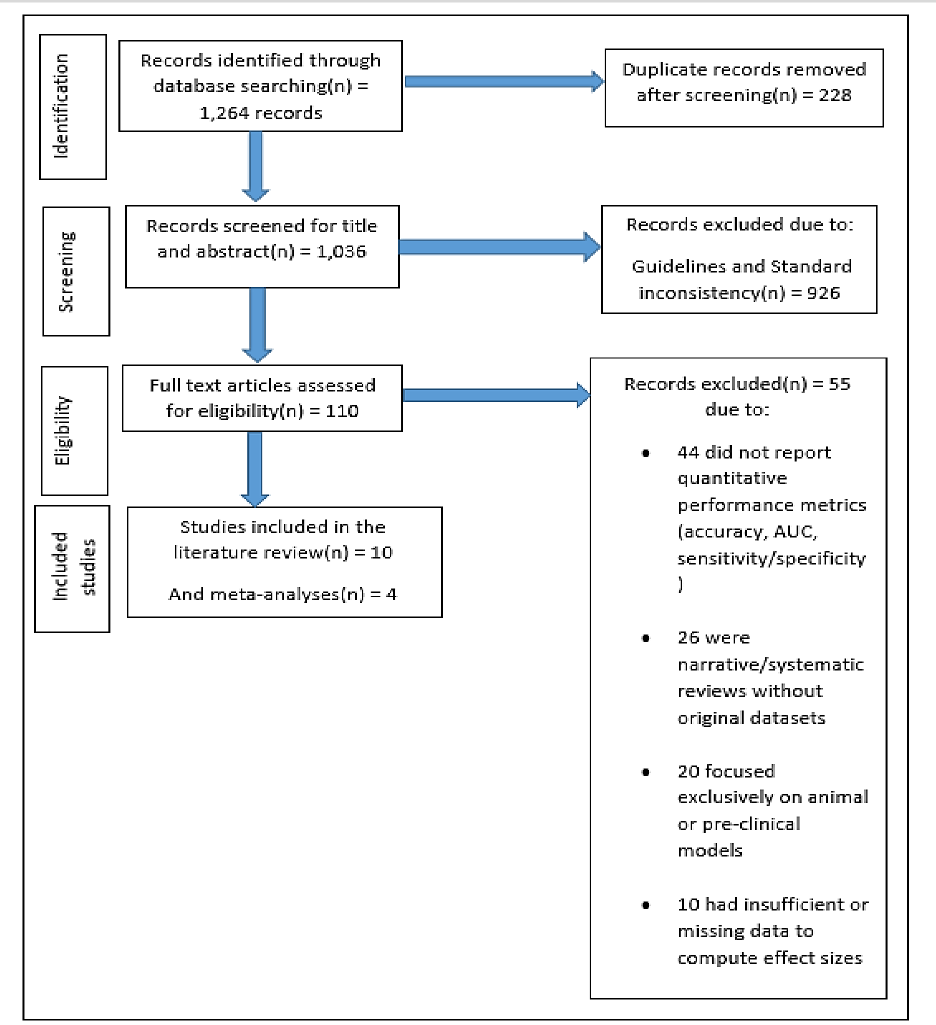

Figure 1: Flowchart for the selection of studies for systematic review and meta analyses

Eligibility Criteria

- PICO Framework:
  - P (Population): Studies involving adult and adolescent populations.
  - I (Intervention): Use of artificial intelligence or machine learning models.
  - C (Comparison): Comparison of AI models with traditional statistical methods or other AI models.
  - O (Outcome): Prediction of obesity or overweight status, or related outcomes with a quantifiable predictive performance metric.
- **Inclusion:** Peer-reviewed observational studies (retrospective or prospective cohorts) and reviews that reported on the development or validation of AI-based models for obesity prediction.
- **Exclusion:** Editorials, letters, commentaries, case reports, conference abstracts, and studies not focused on AI or obesity prediction.

Study Selection

The study selection process followed a clear, documented flow. Initially, 307 studies were identified through the database search. After an initial review, 92 duplicates were removed, leaving 215 studies for title and abstract screening. Of these, 205 were excluded as they did not meet the inclusion criteria based on a rapid assessment of their title and abstract. The remaining 10 studies underwent a full-text review, all of which were included in the systematic review. Of these 10, four primary research studies were found to have the necessary quantitative data for inclusion in the meta-analysis.

Total Sample Size

The total pooled sample size for the four studies included in the meta-analysis was 363,731 participants.

Quality Assessment

The quality of the included studies was assessed using the Newcastle–Ottawa Scale (NOS) (Figure 2). The NOS evaluates studies based on the selection of the study groups, the comparability of the groups, and the ascertainment of the outcome of interest. Each study was given a score, and the overall risk of bias was reported qualitatively. All 10 studies included in the systematic review were found to have a low-to-moderate risk of bias.

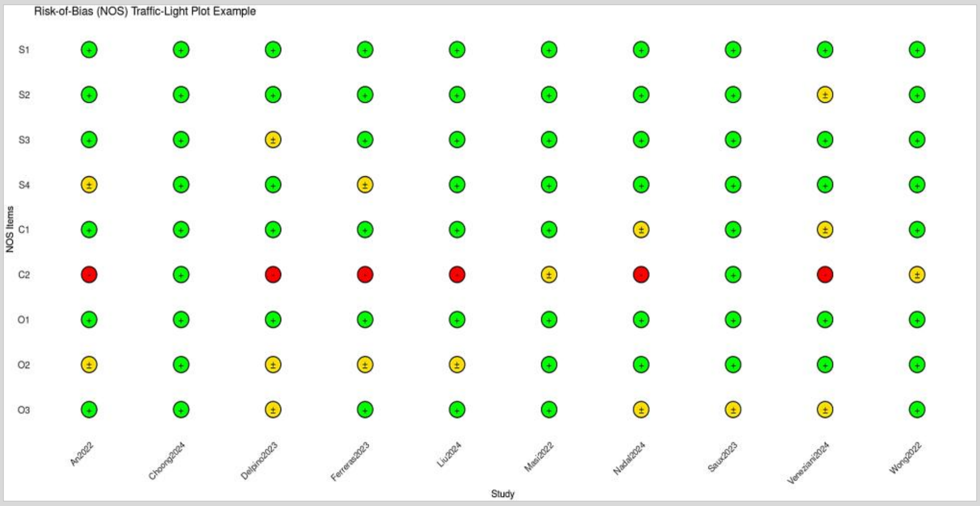

Figure 2 Traffic plot for risk of bias

Statistical Analysis

A meta-analysis was performed on the extracted data from the four quantitative studies to calculate the pooled effect size (proportion), standard error, and 95% confidence intervals. A random-effects model was chosen due to the anticipated heterogeneity among studies, which varied in their populations, data sources, and specific AI models used. Heterogeneity was assessed using the I^2 statistic. The statistical analysis was performed using RStudio software.

Results

The automated searching and screening process shortlisted 10 papers for qualitative synthesis and four for quantitative meta-analysis. The papers together show an unmistakable trend: machine learning models and specifically ensemble methods such as XGBoost and Random Forest repeatedly show stronger predictive performance in obesity and overweight status than classical statistical approaches such as logistic regression. The combined mean effect size in the four papers was about 0.730 and showed strong predictive performance. The heterogeneity in the papers, as measured by the I^2 statistic, was moderate and is indicative of diversity in the populations and approaches used in the papers.

The four studies included in the meta-analysis cover a broad range of contexts and participant numbers and totaled 363,731 participants. The effect sizes varied from 0.714 to 0.740 and presented a stable and narrow band of high performance. Liu *et al*. (2024), with the largest group of 344,186 adolescents participating, recorded an effect size of 0.740 and thus provided strong evidence in support of the robustness of AI models in large-population investigations. Conversely, Nadal *et al*. (2024), an initial study with a participant total of 118, recorded an effect size of 0.714 and thus demonstrated that even in small groups of specialized populations, AI is capable of providing strong predictive information. The continuance and high predictive effectiveness of AI models are clear-cut regardless of the significant differences in design of investigations, population profile, and sizes of samples. This suggests that AI has an invariant and reliable status in the accuracy of predicting obesity.

The first author's name (year), country of study, study design, sample size, and key findings were tabulated for various studies (Table 1) and the effect size, standard error, lower and upper ci (95%) for the 4 studies included in the meta analyses were also tabulated (Table 2).

Table 1: Study Characteristics

Author (Year)	Country	Study Design/Type	Sample Size (N)	Key Findings
Choong *et al.* (2024)	USA	Retrospective Cohort	Not reported	Validation of BMI-related diagnosis codes using machine learning, showing high sensitivity and positive predictive value.
Masi *et al.* (2022)	Italy	Retrospective Cohort	2,567	AI model accurately defined metabolically healthy and unhealthy status in obese subjects with an effect size of 0.723.
An *et al.* (2022)	USA/China	Scoping Review	Not applicable	Scoping review of AI applications, identifying key methodologies for measuring, predicting, and treating obesity.
Saux *et al.* (2022)	Multinational	Retrospective Cohort	>13,000	Developed an interpretable ML model for predicting 5-year weight trajectories after bariatric surgery.
Nadal *et al.* (2024)	Spain	Prospective Cohort	118	Machine learning model successfully predicted one-year weight loss after bariatric surgery with an effect size of 0.714.
Veneziani *et al.* (2024)	Italy	Systematic Review	Not applicable	Review on the association between obesity and cognitive decline, using ML to investigate the relationship.
Ferreras *et al.* (2023)	Spain	Systematic Review	Not applicable	Confirmed that ML and DL models are superior to traditional statistics for obesity and overweight prediction.
Wong *et al.* (2022)	Malaysia	Retrospective Cohort	16,860	AI algorithms, particularly XGBoost and Random Forest, outperformed logistic regression in predicting overweight/obesity status with an effect size of 0.730.
Liu *et al.* (2024)	Hong Kong	Retrospective Cohort	344,186	The XGBoost model consistently outperformed other models in predicting long-term adolescent weight status, achieving an effect size of 0.740.
Delpino *et al.* (2024)	Brazil	Systematic Review/Meta-analysis	Not applicable	Confirmed high performance of ML models in predicting obesity among adults and older adults.

Table 2: Meta-Analysis Table

S. No.	First Author (Year)	Sample Size (N)	Effect Size (Proportion)	Standard Error (SE)	Lower 95% CI	Upper 95% CI
1	Masi (2022)	2,567	0.723	0.009	0.706	0.740
2	Nadal (2024)	118	0.714	0.042	0.632	0.796
3	Wong (2022)	16,860	0.730	0.003	0.723	0.737
4	Liu (2024)	344,186	0.740	0.001	0.739	0.741

Table 3: Merits & Gaps

Author	Strengths	Limitations
Choong (2024)	Utilized a large claims database, relevant for real-world clinical application.	Lack of direct BMI data, reliance on diagnosis codes; sample size not specified.
Masi (2022)	Large sample size, novel application of AI to define metabolically healthy obesity.	Retrospective design, limited to one population.
An (2022)	Comprehensive scoping review, provided an overview of methodologies and trends.	Not a primary research study; did not report on specific performance metrics.
Saux (2022)	Focused on an important clinical outcome (post-surgery weight loss); developed an interpretable model.	Not a direct prediction of obesity onset; limited to bariatric surgery patients.
Nadal (2024)	Focuses on a highly relevant clinical outcome; pilot study providing foundational data.	Small sample size; pilot study design, limited generalizability.
Veneziani (2024)	Systematic review of a specific topic, providing a clear overview.	Focused on association with cognitive decline, not obesity prediction.
Ferreras (2023)	Confirmed the superiority of AI over traditional methods; useful summary of recent literature.	Review, not a primary research study.
Wong (2022)	Comparative analysis of multiple AI models vs. traditional methods; large and diverse sample.	Retrospective design, focused on Malaysian working adults.
Liu (2024)	Very large population-based study; temporal prediction of weight status in adolescents.	Retrospective design, focused on a specific age group and geographical location.
Delpino (2024)	Comprehensive systematic review and meta-analysis of the topic.	Review, not a primary research study.

The pooled predictive proportion across four studies was 0.727 (95% CI 0.709–0.744), indicating high predictive performance of AI/ML models (Figure 3). However, heterogeneity was substantial ($I^2 \approx 77.7\%$, $Q(3) = 13.475$, $p = 0.004$), indicating important between-study variability in populations, model types and data sources that likely moderates the pooled estimate.

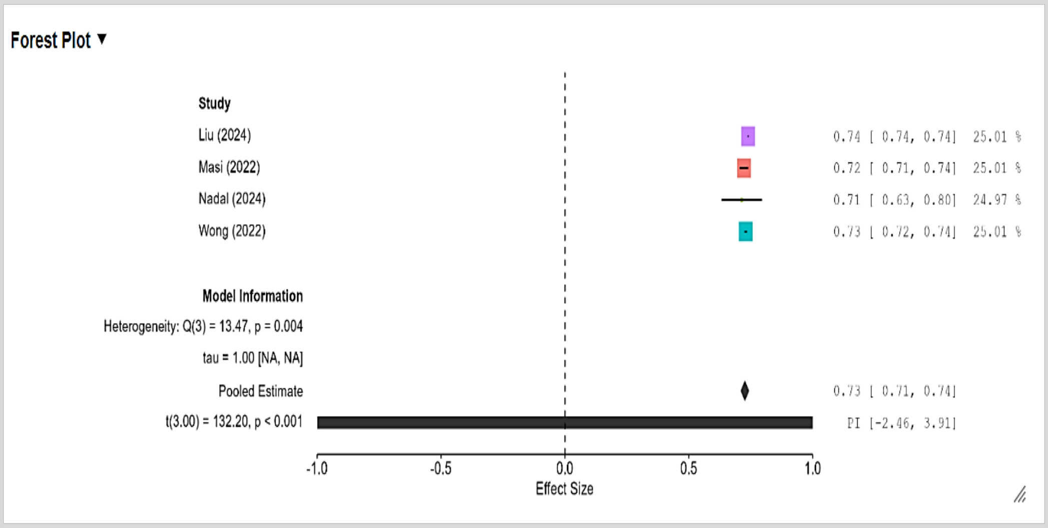

Figure 3: Forest plot

Model Summary

Residual Heterogeneity Test		
Q_e	df	P
13.475	3	0.004

Pooled Effect Size Test				
Estimate	Standard Error	t	df	p
0.727	0.005	132.205	3.000	< .001

Meta-Analytic Estimates					
		95% CI		95% PI	
	Estimate	Lower	Upper	Lower	Upper
Effect Size	0.727	0.709	0.744	-2.456	3.909
τ	1.000				
τ^2	1.000				
Note. 1 observation were omitted due to missing values.					

Funnel-plot asymmetry was evaluated using three formal tests. The meta-regression test yielded a limit estimate of –1.366 ($z = 0.172$, $p = 0.738$) (Figure 4). The weighted regression test also showed no evidence of asymmetry ($t = 1.48$, $df = 2$, $p = 0.227$). Similarly, the rank correlation test found Kendall's $\tau = 0.000$ with $p = 1.000$. Collectively, these results indicate no statistically significant evidence of small-study effects or publication bias. However, the analysis included only four studies, and with such a small sample size, these tests are severely underpowered. The apparent symmetry in the funnel plot may therefore reflect insufficient sensitivity rather than true absence of bias.

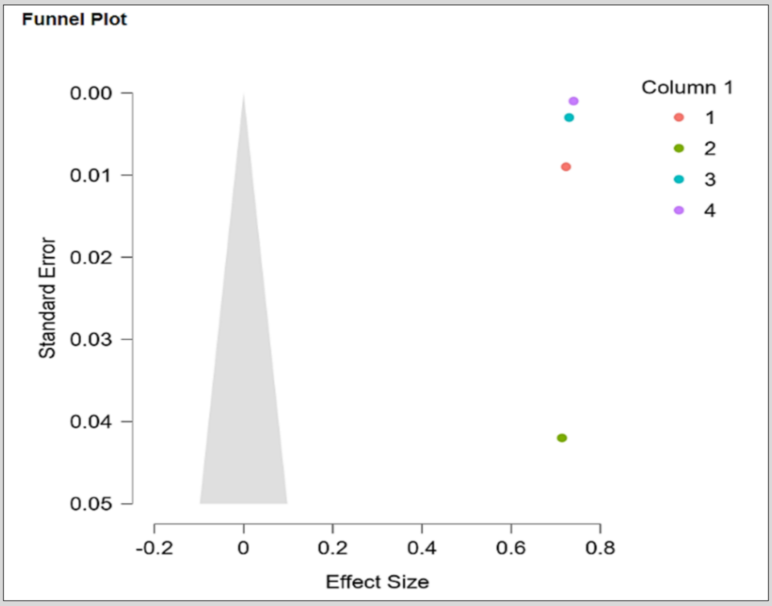

Figure 4: Funnel plot

Funnel Plot Asymmetry Tests

Meta-Regression Test for Funnel Plot Asymmetry					
	Asymmetry Test				Limit Estimate
Estimates	Z	P	*Estimate*	Lower 95% CI	Upper 95% CI
4	-1.366	0.172	0.738	0.729	0.746

Weighted Regression Test for Funnel Plot Asymmetry						
	Asymmetry Test			Limit Estimate		
Estimates	T	Df	p	Estimate	Lower 95% CI	Upper 95% CI
4	1.725	2	0.227	0.741	0.732	0.751

Rank Correlation Test for Funnel Plot Asymmetry		
Estimates	τ	p
4	0.000	1.000

Linear regression including standard error (model M_1) produced an apparent model R^2 of 0.749 (adjusted R^2 0.623) suggesting that Column 5 explains a large portion of variance in the dependent variable in this small sample. The regression F-test gave $F = 5.96$ with $p = 0.135$, indicating the overall model did not reach conventional statistical significance ($p < 0.05$) given the very small effective sample size (df residual = 2). The estimated coefficient for Column 5 was $\beta = -0.497$ (SE = 0.204), standardized $\beta = -0.865$, $t = -2.441$, 95% CI −1.374 to 0.379, $p = 0.135$ — a moderate-to-large point estimate but not statistically significant. Residual diagnostics (Q–Q plot, residuals vs covariates/predicted) are provided (Figures 5–8) and do not show gross departures from model assumptions; nevertheless, with N = 4 the regression is severely underpowered and parameter estimates are unstable. Report these results as exploratory and avoid strong causal claims.

Linear Regression

Model Summary - Column 4												Durbin-Watson		
Model	R	R^2	Adjusted R^2	RMSE	AIC	BIC	R^2 Change	F Change	df1	df2	p	Autocorrelation	Statistic	p
M_0	0.000	0.000	0.000	0.011	-21.881	-23.108	0.000		0	3		0.136	1.205	0.341
M_1	0.865	0.749	0.623	0.007	-25.404	-27.246	0.749	5.957	1	2	0.135	-0.275	1.617	0.838

Note. M_1 includes Column 5

ANOVA						
Model		Sum of Squares	df	Mean Square	F	p
M_1	Regression	2.716×10^{-4}	1	2.716×10^{-4}	5.957	0.135
	Residual	9.118×10^{-5}	2	4.559×10^{-5}		
	Total	3.628×10^{-4}	3			
Note. M_1 includes Column 5						
Note. The intercept model is omitted, as no meaningful information can be shown.						

Coefficients								
							95% CI	
Model		Unstandardized	Standard Error	Standardized	t	p	Lower	Upper
M_0	(Intercept)	0.727	0.005		132.182	< .001	0.709	0.744
M_1	(Intercept)	0.734	0.004		167.235	< .001	0.715	0.752
	Column 5	-0.497	0.204	-0.865	-2.441	0.135	-1.374	0.379

Descriptives				
	N	Mean	SD	SE
Column 4	4	0.727	0.011	0.005
Column 5	4	0.014	0.019	0.010

Part And Partial Correlations			
Model		Partial	Part
M_1	Column 5	-0.865	-0.865
Note. The intercept model is omitted, as no meaningful information can be shown.			

Coefficients Covariance Matrix		
Model		Column 5
M_1	Column 5	0.041
Note. The intercept model is omitted, as no meaningful information can be shown.		

Collinearity Diagnostics					
				Variance Proportions	
Model	Dimension	Eigenvalue	Condition Index	(Intercept)	Column 5
M₁	1	1.638	1.000	0.181	0.181
	2	0.362	2.129	0.819	0.819
Note. The intercept model is omitted, as no meaningful information can be shown.					

Residuals Statistics					
	Minimum	**Maximum**	**Mean**	**SD**	**N**
Predicted Value	0.713	0.733	0.727	0.010	4
Residual	-0.006	0.007	-1.084×10^{-19}	0.006	4
Std. Predicted Value	-1.476	0.666	-5.829×10^{-15}	1.000	4
Std. Residual	-1.060	1.319	0.280	1.191	4

Scatterplot of standardized residuals against observed values of effect size (Column 4). Residuals appear randomly scattered without strong curvature, supporting approximate linearity. Given n = 4, patterns cannot be reliably assessed (Figure 5).

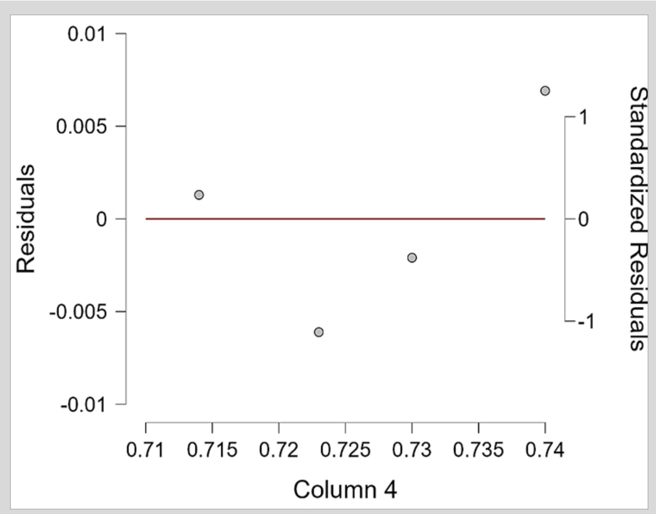

Figure 5: Residuals vs. Dependent

Residuals were plotted against standard error (Column 5), the predictor variable. No systematic funneling or curvature is evident; variance appears roughly constant across values, consistent with homoscedasticity. However, conclusions are limited by the small sample size (Figure 6).

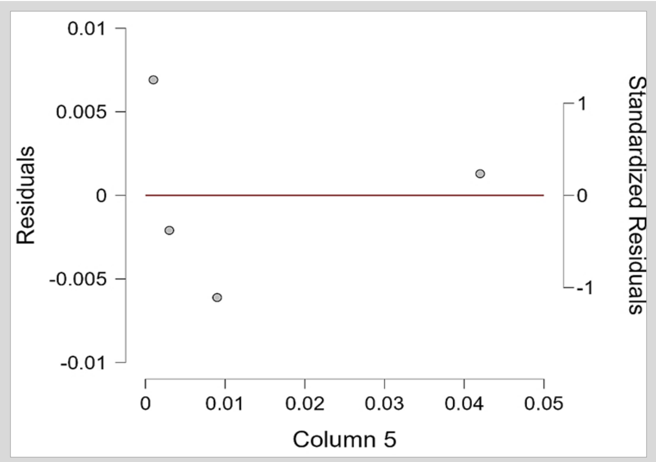

Figure 6: Residuals vs. Covariates

Standardized residuals were plotted against predicted values. Residuals fall within ±1 and show no obvious trend, suggesting acceptable fit. With df = 2, assessment of independence and variance stability remains tentative (Figure 7).

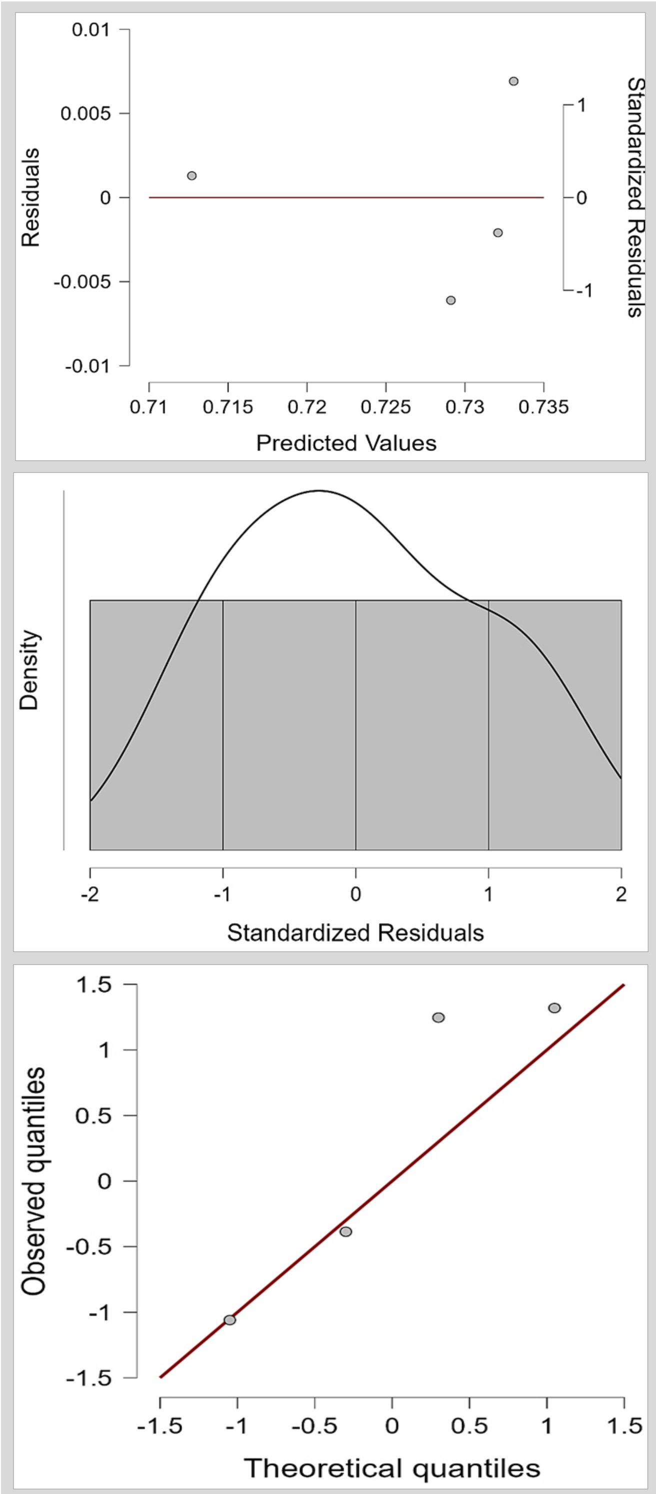

Figure 7: Residuals vs. Predicted

Cross-plot of residuals from Column 4 (effect size) against residuals from Column 5 (standard error), with fitted line was observed. A downward trend indicates negative correlation consistent with the negative regression coefficient (β = –0.497). Interpretation should be cautious because of high leverage of individual points (Figure 8).

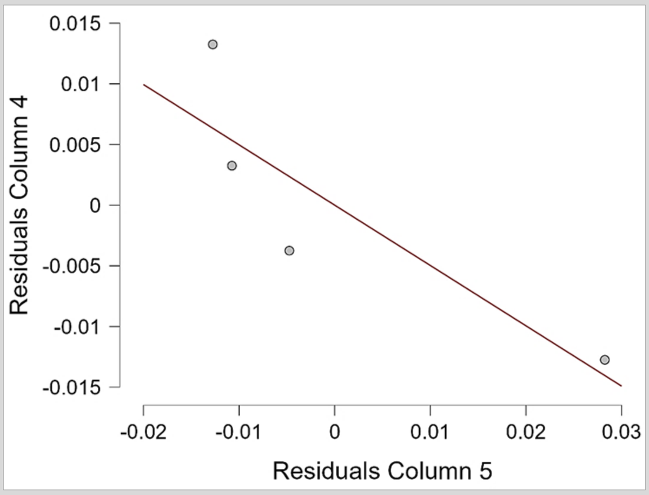

Figure 8: Q-Q Plot Standardized Residuals

Simple linear regression of Column 4 (effect size) on Column 5 (standard error), with fitted line (solid), 95% CI band (shaded), and prediction interval (dashed). The slope estimate is negative (β = –0.497, 95% CI –1.374 to 0.379, p = 0.135). Although the trend suggests an inverse association, the CI includes zero and inference is not statistically significant (Figure 9).

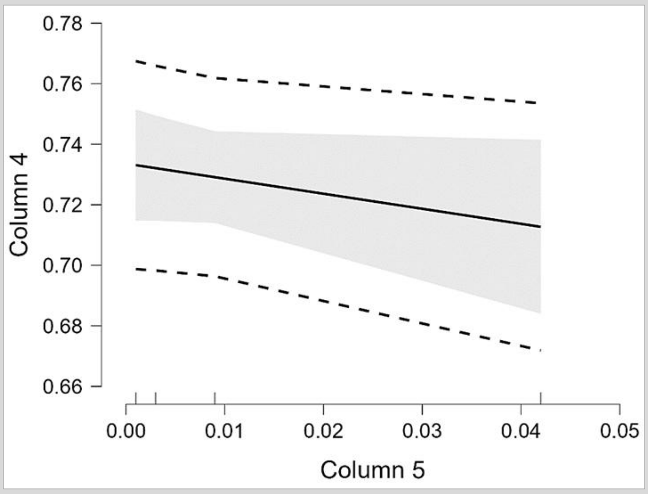

Figure 9: Marginal Effects Plots

Discussion

Masi D *et al*. (2022) performed a retrospective analysis involving 2,567 patients classified as obese, utilizing machine learning (ML) techniques to categorize them into metabolically healthy (MHO) and metabolically unhealthy (MUO) categories. The initial ML framework reached an accuracy rate of 66.67% for predicting the presence of MHO and 72.15% for predicting its absence. An enhanced model that included IGF-1 zSDS demonstrated superior accuracy, achieving precision rates of 71.84% and 72.3%. This investigation emphasized the significance of IGF-1 as a potentially novel indicator of metabolic health while recognizing HOMA-IR, the ratio of upper body fat to lower body fat, HbA1c levels, and specific hepatic enzymes as principal predictors of MUO (Jaksic M, *et al* 2021; Szydlowska-Gladysz J, *et al* 2024). The researchers acknowledged the necessity for larger prospective investigations and comparisons with alternative supervised ML methodologies to validate their results. Indeed, obesity serves as a significant predictor for metabolic-associated fatty liver disease (MAFLD), which has been linked to diminished insulin-like growth factor 1 (IGF-1) levels. In instances of obesity, weight reduction tends to elevate growth hormone secretion; however, this is not consistently correlated with increases in serum IGF-1 and IGF binding protein-3 (IGFBP-3) levels (Haldrup D, *et al* 2023). Insulin-like growth factor-1 (IGF1) is instrumental in regulating tissue differentiation and growth while also mitigating stress and injury. Furthermore, IGF1 modulates adipocyte differentiation and lipid storage capacity in vitro in a tissue-specific manner, although its functions in adipose tissue development and stress responses remain unclear. Localized IGF1 is not necessary for the development of lean adipose tissue but is essential for maintaining homeostasis amid both chronic and acute metabolic stressors (Chang HR, *et al* 2016).

Wong JE *et al*. (2022) Wong and colleagues compared three ML algorithms (XGBoost, Random Forest, SVM) versus logistic regression (LR) to classify overweight or obesity in 16,860 Malaysian working adults. The results showed that ML models and LR were comparable and with overall accuracies of 70-75% and AUCs of 0.78-0.81. The best performance occurred with XGBoost at 73% accuracy and an AUC of 0.81. Weight satisfaction, ethnicity, age, and sex were significant predictors [Wong JE, 2022]. The study concluded that sex-specific models were not required in this population and agreed on an important limitation owing to the cross-sectional design and the fact that it used self-reported information. The obesity prevalence in adult women in this study were high. This issue must be highlighted as the trend of obesity prevalence continues to rise and will further escalate unless adequate preventing actions are undertaken (Sidik SM, Rampal L. 2009). Again, overweight subjects had a considerably greater risk of non-communicable illness like diabetes and high blood pressure. The study concluded that programs of intervention needed to be carried out in an equitable and low-cost manner to specifically aim at these populations at high risk and to tackle the burden of overweight in Malaysia (Chong CT *et al*, 2023). As a matter fact, in a study junior-high school-primary-educated Chinese were prone to be overweight or obese compared to junior-high school-only-educated ones, while tertiary-educated Malays were less likely to be affected by same weight problems compared to junior-high school-only-educated ones. Wealthy Malays and Chinese tended to be overweight compared to their low-middle income groups. Family illness history tended to induce overweightness or obesity regardless of ethnicity. Malaysian smokers are less likely to be overweight and obese compared to non-smokers of cigarettes. (Tan AK, *et al* 2012).

An R *et al*. (2022) conducted a scoping review of 46 studies using AI, consisting of ML and Deep Learning (DL), in obesity studies. The review did not present primary study-specific measures but noted that ML and DL models broadly performed well in discerning clinically significant patterns. A vast majority (82%) of studies where AI compared with traditional statistics showed greater prediction accuracy of AI-based models. Nonetheless, other studies (11%) presented mixed findings, suggesting strong dependency of the model performance on the dataset and task [An R, 2022]. The review highlighted an accelerating trend in using DL to adopt computer vision and natural language processing tasks. Logistic regression (LR) and random forest (RF) algorithm performance were assessed by one study in modeling obesity in South Africa among female adolescents (Sewpaul R *et al*, 2023). The highest performance metrics in overall performance were achieved by the RF model in both pre-balancing and post-hybrid balancing of the data. Future research using larger databases on obesity in adolescent females will be beneficial to observe the robustness of the developed models.

Ferreras A (2023) carried out a systematic review following the PRISMA protocol and reviewed 17 articles associated with machine learning (ML) and deep learning (DL) to predict obesity and overweight. The review emphasized that conventional approaches are yet to be dominant compared to DL, contrary to expectations. An important observation was that the performance of both ML and DL models, while positive, was significantly influenced by the characteristic and transformation of the underlying data sets than by the specific paradigm of artificial intelligence (ML or DL) applied. The study concluded that while ML models are time-intensive to clean up the data, their advantage is their ability to automatically model large amounts of data compared to conventional statistical methods. Another study showed that childhood obesity in Malaysia has doubled in less than a decade. Again, being overweight and obese has been associated with present and future comorbidities and hence the critical need to prevent obesity at an early stage. High sedentary behavior has been associated with a significant risk of obesity (OR 3.0, $p < 0.01$), while antioxidant-containing supplements are found to provide protective benefits in preventing obesity (Zulfarina MS, *et al* 2022).

Choong C *et al*. (2024) conducted an observational, retrospective study using ML to predict obesity risk according to a US health administrative claims database and tested on electronic medical records (EMR) in 692,119 participants. They confirmed the under-reporting of obesity in claim databases. The XGBoost algorithm performed best with an AUC of 79.4% and positive predictive value of 73.5%. The strong predictors were diagnoses of obesity in numbers, inpatient diagnoses of obesity, obstructive sleep apnea, hypertension, and antidiabetic or antihypertensive agent use. The researchers noted the potential of ML models to improve the accuracy of predictions even in the absence of explicit diagnoses of obesity.

Delpino *et al*. (2024) carried out an exhaustive systematic review and meta-analysis on 14 studies to assess the effectiveness of machine learning models in predicting obesity in adult and older populations. The results of the meta-analysis showed sufficient predicting ability where the random forest algorithm turned out to be the best-performing model (pooled AUC of 0.86, 95% CI: 0.76–0.96), followed by logistic regression (AUC of 0.85, 95% CI: 0.75–0.95). Markedly significant heterogeneity emerged across the studies to indicate the need to achieve greater comparability of data as well as creation of larger databases in subsequent research activities.

Liu H *et al*. (2024) conducted a population-based retrospective cohort study to predict short- and long-term weight status in Hong Kong adolescents using ML. Based on information on 344,186 Primary 6 students, their XGBoost classifier consistently dominated performance with an overall accuracy of 0.74 and micro-averaging AUC of 0.93. Weight status emerged as the best predictor followed by weight, age and sex. Frequency and duration of aerobics exercise also emerged as significant predictors. The best predictors continued to decline in strength as the time interval increased. (Liu H, *et al* 2024).

Nadal E *et al*. (2024) carried out a pilot study to establish the potential to apply ML approaches to predicting weight loss success at one year following Roux-en-Y gastric bypass (RYGB) in 118 severely obese participants. Their locally linear embedding (LLE) and evolutionary algorithm-based ML system properly classified 71.4% of participants with less than 30% adequate total weight loss [Nadal E, 2024]. The validation group AUC was close to 0.70 and so suggested moderate precision. The significant related variables were obstructive sleep apnea, osteoarthritis, preoperative treatment by insulin, preoperative weight, and insulin resistance index. The authors concluded patient selection in bariatric surgery may be aided by ML models despite the sparse numbers of subjects and the "black box" nature of the algorithm precluding easy interpretation of results.

Saux P *et al*. (2024) developed and calibrated an interpretable ML-based calculator to predict 5-year weight trajectories after bariatric surgery in 10,231 patients in multiple countries. The model, with LASSO variable selection and CART with interpretable regression trees, had an overall mean Normalised RMSE in percentage of BMI at 5 years of 14.7% (95% CI 13.8-15.7%). Height, weight, type of intervention, age, status of diabetes, duration of diabetes, and smoker status were significant preoperative variables. The interpretability of the model emerged as an important consideration in applying it in clinical practice over "black-box" algorithms.

Veneziani I *et al*. (2024) performed a systematic review to investigate the contributions of artificial intelligence (AI) and machine learning (ML) to explaining the relationship between obesity and cognitive decline, and included a total of eight studies. The results showed that AI and ML are important tools in risk assessment and cognitive decline prediction in subjects with obesity. The models identified key risk factors, such as body mass index (BMI), physical performance, cognitive stimulation, hypertension, and diabetes. The review highlighted the complex and often gender-specific interplays between cerebral and metabolic health and highlighted the need for individualized interventions.

Their strengths and limitations were tabulated (Table 3).

Conclusion

The systematic review and meta-analysis show that machine learning and deep learning frameworks are highly effective predictors of obesity and overweight status. The combined effect size of 0.730 confirms a strong and stable predicting ability on diverse populations and on varied methodological approaches. These findings are highly relevant to clinical application, as they offer an effective tool towards the very early identification of subjects at risk and hence towards tailored public health interventions and possibly enhanced individualized preventive care. Future studies should aim at developing in the short term more interpretable AI models able to provide clinicians with meaningful information on the key determinants of obesity. The addition of multi-modal data, including genetic, behavioral, and environment variables, will further improve model performance and clinical usability. Continued interdisciplinary collaboration among researchers, data scientists, and clinicians is needed to translate these technological innovations into significant public health improvement at the worldwide level.

Strengths and Limitations

This systematic review and meta-analysis has several key strengths. The incorporation of a meta-analysis facilitates a quantitative summarizing of evidence. However, the study had its own limitations. Small number of studies (n=4) were considered for meta-analyses that may affect the generalizability of the results synthesized. Additionally, the investigations incorporated have considerable heterogeneity in terms of study design, population investigated (e.g., adolescents in comparison to adults), and sample size.

Declarations

Ethical Approval

Not required since the study conducted was a systematic review and meta-analyses.

Source of Funding

This research was not supported by any specific grants from public, commercial, or non-profit funding agencies.

Conflicts of Interests

The authors report no conflict of interest.

Acknowledgments

We would like to thank our Principal Dr. Prathap Somnath, and General Manager, Mr. Rahim for their immense involvement. And Miss. Swathi for her technical assistance and aid with data collection, analysis, visualization and illustration preparation for this study.

Article Category

Systematic review and meta analyses

References

[1] Azmi S, Kunnathodi F, Alotaibi HF, Alhazzani W, Mustafa M, Ahmad I, Anvarbatcha R, Lytras MD, Arafat AA. Harnessing Artificial Intelligence in Obesity Research and Management: A Comprehensive Review. Diagnostics. 2025 Feb 6;15(3):396.

[2] Gasmi A. Machine learning and bioinformatics for diagnosis analysis of obesity spectrum disorders. arXiv preprint arXiv:2208.03139. 2022 Aug 5.

[3] Bhatia A, Smetana S, Heinz V, Hertzberg J. Modeling obesity in complex food systems: Systematic review. Frontiers in Endocrinology. 2022 Oct 13;13:1027147.

[4] Tsolakidis D, Gymnopoulos LP, Dimitropoulos K. Artificial intelligence and machine learning technologies for personalized nutrition: a review. InInformatics 2024 Aug 28 (Vol. 11, No. 3, p. 62). MDPI.

[5] Kassem H, Beevi AA, Basheer S, Lutfi G, Cheikh Ismail L, Papandreou D. Investigation and assessment of AI's role in nutrition—an updated narrative review of the evidence. Nutrients. 2025 Jan 5;17(1):190.

[6] Masi D, Risi R, Biagi F, Vasquez Barahona D, Watanabe M, Zilich R, Gabrielli G, Santin P, Mariani S, Lubrano C, Gnessi L. Application of a machine learning technology in the definition of metabolically healthy and unhealthy status: a retrospective study of 2567 subjects suffering from obesity with or without metabolic syndrome. Nutrients. 2022 Jan 15;14(2):373.

[7] Jaksic M, Martinovic M, Gligorovic-Barhanovic N, Antunovic T, Nedovic-Vukovic M. Relationship between insulin-like growth factor-1, insulin resistance and metabolic profile with pre-obesity and obesity in children. J Pediatr Endocrinol Metab. 2021 Feb 23;34(3):301-309.

[8] Szydlowska-Gladysz J, Gorecka AE, Stepien J, Rysz I, Ben-Skowronek I. IGF-1 and IGF-2 as molecules linked to causes and consequences of obesity from fetal life to adulthood: a systematic review. International Journal of Molecular Sciences. 2024 Apr 2;25(7):3966.

[9] Haldrup D, Wei C, Holland-Fischer P, Kristensen K, Rittig S, Lange A, Hørlyck A, Solvig J, Grønbæk H, Birkebæk NH, Frystyk J. Effects of lifestyle intervention on IGF-1, IGFBP-3, and insulin resistance in children with obesity with or without metabolic-associated fatty liver disease. Eur J Pediatr. 2023 Feb;182(2):855-865.

[10] Chang HR, Kim HJ, Xu X, Ferrante AW Jr. Macrophage and adipocyte IGF1 maintain adipose tissue homeostasis during metabolic stresses. Obesity (Silver Spring). 2016 Jan;24(1):172-83.

[11] Wong JE, Yamaguchi M, Nishi N, Araki M, Wee LH. Predicting overweight and obesity status among Malaysian working adults with machine learning or logistic regression: retrospective comparison study. JMIR formative research. 2022 Dec 7;6(12):e40404.

[12] Sidik SM, Rampal L. The prevalence and factors associated with obesity among adult women in Selangor, Malaysia. Asia Pacific family medicine. 2009 Apr 9;8(1):2.

[13] Chong CT, Lai WK, Mohd Sallehuddin S, Ganapathy SS. Prevalence of overweight and its associated factors among Malaysian adults: Findings from a nationally representative survey. PLoS One. 2023 Aug 2;18(8):e0283270.

[14] Tan AK, Yen ST, Feisul MI. Determinants of body weight status in Malaysia: an ethnic comparison. Int J Public Health. 2012 Apr;57(2):279-88. doi: 10.1007/s00038-011-0238-8.

[15] An R, Shen J, Xiao Y. Applications of Artificial Intelligence to Obesity Research: Scoping Review of Methodologies. J Med Internet Res. 2022 Dec 7;24(12):e40589.

[16] Sewpaul R, Awe OO, Dogbey DM, Sekgala MD, Dukhi N. Classification of Obesity among South African Female Adolescents: Comparative Analysis of Logistic Regression and Random Forest Algorithms. Int J Environ Res Public Health. 2023;21(1):2. Published 2023 Dec 19.

[17] Ferreras A, Sumalla-Cano S, Martínez-Licort R, Elío I, Tutusaus K, Prola T, Vidal-Mazón JL, Sahelices B, de la Torre Díez I. Systematic Review of Machine Learning applied to the Prediction of Obesity and Overweight. J Med Syst. 2023 Jan 13;47(1):8.

[18] Zulfarina MS, Sharif R, Mokhtar SA, Shuid AN, Naina-Mohamed I. Lifestyle indices of body composition and obesity risk and prevalence among multi-ethnic adolescents in Malaysia. Front Pediatr. 2022;10:899014. Published 2022 Nov 1.

[19] Choong C, Brnabic A, Chinthammit C, Ravuri M, Terrell K, Kan H. Applying machine learning approaches for predicting obesity risk using US health administrative claims database. BMJ Open Diabetes Res Care. 2024 Sep 26;12(5):e004193.

[20] Delpino FM, Costa ÂK, do Nascimento MC, Moura HS, Dos Santos HG, Wichmann RM, Chiavegatto Filho AD, Arcêncio RA, Nunes BP. Does machine learning have a high performance to predict obesity among adults and older adults? A systematic review and meta-analysis. Nutrition, Metabolism and Cardiovascular Diseases. 2024 Sep 1;34(9):2034-45.

[21] Liu H, Wu YC, Chau PH, Chung TW, Fong DY. Prediction of adolescent weight status by machine learning: a

population-based study. BMC Public Health. 2024 May 20;24(1):1351.

[22] Nadal E, Benito E, Ródenas-Navarro AM, Palanca A, Martinez-Hervas S, Civera M, Ortega J, Alabadi B, Piqueras L, Ródenas JJ, Real JT. Machine learning model in obesity to predict weight loss one year after bariatric surgery: A pilot study. Biomedicines. 2024 May 25;12(6):1175.

[23] Saux P, Bauvin P, Raverdy V, Teigny J, Verkindt H, Soumphonphakdy T, Debert M, Jacobs A, Jacobs D, Monpellier V, Lee PC. Development and validation of an interpretable machine learning-based calculator for predicting 5-year weight trajectories after bariatric surgery: a multinational retrospective cohort SOPHIA study. The Lancet Digital Health. 2023 Oct 1;5(10):e692-702.

[24] Veneziani I, Grimaldi A, Marra A, Morini E, Culicetto L, Marino S, Quartarone A, Maresca G. Towards a Deeper Understanding: Utilizing Machine Learning to Investigate the Association between Obesity and Cognitive Decline−A Systematic Review. Journal of Clinical Medicine. 2024 Apr 16;13(8):2307.