IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Ingénierie des Systèmes d'Information
Vol. 31, No. 2, February, 2026, pp. 539-549
Journal homepage: http://iieta.org/journals/isi

Explainable Machine Learning for Early Malnutrition Risk Prediction Using Clinical and Socioeconomic Determinants

Dino Quinteros-Navarro[1*], Melissa Soto[2], Diana Quispe-Arbildo[2,3], Diana Ponce[2,4], Carmen Marquez[2], Miriam Rojas[2], Ana Maguiña[2]

[1] Facultad de Ingeniería, Carrera Profesional de Ingeniería de Sistemas, Universidad Tecnológica del Perú, Lima 15046, Perú
[2] Faculty of Medicine, Universidad Nacional Federico Villarreal, Lima 15007, Perú
[3] Unidad de Soporte Nutricional Artificial, Hospital Alberto Sabogal Sologuren, Callao 07011, Perú
[4] Unidad de Cuidados Intensivos, Hospital Nacional Dos de Mayo, Lima 15001, Perú

Corresponding Author Email: C29925@utp.edu.pe

Copyright: ©2026 The authors. This article is published by IIETA and is licensed under the CC BY 4.0 license (http://creativecommons.org/licenses/by/4.0/).

https://doi.org/10.18280/isi.310222

Received: 16 September 2025
Revised: 1 December 2025
Accepted: 19 February 2026
Available online: 28 February 2026

Keywords:
malnutrition risk prediction, machine learning, XGBoost, explainable artificial intelligence, socioeconomic determinants

ABSTRACT

Early identification of malnutrition risk remains a major challenge in heterogeneous clinical settings, where conventional screening tools may have limited predictive accuracy and adaptability. This retrospective proof-of-concept study developed and internally validated machine learning models for malnutrition risk prediction using 1,000 anonymized patient records from a public referral hospital in Lima, Peru. The dataset integrated demographic, clinical, biochemical, and socioeconomic variables, including age, body mass index, serum albumin, food insecurity, and household income. Random Forest, XGBoost, and deep neural network models were evaluated using stratified 10-fold cross-validation and compared with conventional nutritional screening approaches. The deep neural network achieved the highest discriminative performance (AUC = 0.90), whereas XGBoost provided the most balanced overall performance and more stable calibration. Explainability analysis based on SHAP identified serum albumin, body mass index, and socioeconomic indicators as the strongest contributors to risk prediction. These findings indicate that machine learning models can improve early malnutrition risk stratification by integrating multidimensional determinants beyond conventional screening variables. However, external validation in independent multicenter cohorts is required before clinical deployment.

1. INTRODUCTION

Malnutrition persists as one of the most intricate and enduring challenges within global public health today. It is not limited to traditional undernutrition; rather, it encompasses a "triple burden" consisting of acute and chronic malnutrition, micronutrient deficiencies, and the rising prevalence of overweight and obesity [1]. This complex phenomenon disproportionately impacts according to the predefined age range established in the study protocol vulnerable demographics, including young children, pregnant women, the elderly, and clinical patients, particularly those living in adverse socioeconomic conditions. Current global estimates indicate that while over 462 million adults suffer from malnutrition, approximately 1,900 million are classified as overweight or obese [2]. Such figures highlight a significant nutritional paradox in which food insecurity and obesity coexist, a trend frequently observed in low-income urban environments like those in Lima, Peru.

Addressing these overlapping nutritional crises requires a sophisticated understanding of how socioeconomic determinants drive both deficiency and excess. In many developing regions, the transition toward processed diets coupled with limited access to quality healthcare exacerbates these disparities. Consequently, identifying at-risk individuals within the public health system remains a priority for reducing long-term morbidity. While these statistics provide a grim global outlook, they underscore the urgent need for targeted interventions and improved screening tools to protect the most susceptible members of society.

The magnitude of the problem is accentuated in contexts of structural vulnerability. In sub-Saharan Africa, for example, the prevalence of child malnutrition remains at alarming levels, with variations between countries influenced by environmental, conflict, and socioeconomic factors [3]. In Zimbabwe, the authors [4] reported that anemia affects more than 24% of women of reproductive age, identifying determinants such as educational level, household wealth, body mass index (BMI), and modern contraceptive use. In Asia, studies in Bangladesh and China have shown that both undernutrition and overweight in children are influenced by diet, parental factors, and socioeconomic status [1, 2]. In addition, food insecurity resulting from armed conflict or population displacement, such as in Somalia and South Sudan, represents a critical factor that increases the burden of malnutrition in displaced populations [5].

The consequences of malnutrition transcend the individual and generate a systemic impact. In childhood, it is associated with stunted growth, lower school performance, and increased susceptibility to infectious diseases, with effects extending into adulthood [1]. In pregnant women, it increases the risk of obstetric complications, premature birth, and neonatal mortality [6]. In older adults, it is linked to sarcopenia, frailty, and greater functional dependence [7]. In hospitalized and surgical patients, malnutrition constitutes an independent predictor of postoperative complications, prolonged stays, and mortality [8, 9]. These adverse effects make malnutrition not only a clinical challenge, but also a socioeconomic and health equity problem.

Traditionally, instruments such as the *MNA, NRS-2002* and the criteria of the *GLIM* have been used to assess nutritional status [10]. However, although useful, these tools have important limitations: they depend on trained personnel, require time to apply and, in many cases, incorporate subjective information, which decreases their reproducibility and scalability. In addition, their sensitivity and specificity in various contexts are insufficient, especially in heterogeneous populations and in environments with limited resources [11].

In this scenario, artificial intelligence (AI) and machine learning (ML) have positioned themselves as innovative alternatives. These technologies allow processing large volumes of clinical, demographic, socioeconomic, environmental and digital data, identifying non-linear patterns and complex relationships that escape traditional statistical methods [12, 13]. Various studies have shown its usefulness in anticipating nutritional risk and its clinical consequences. For example, the study [14] evaluated the performance of the MUST-Plus model in the Mount Sinai health system (USA), evidencing calibration problems between sex and race that could be corrected through logistic recalibration, highlighting the importance of equity in predictive models. The study [10] developed an AI model based on GLIM criteria applied to patients with acute abdomen, achieving greater diagnostic accuracy than conventional methods. In cardiovascular surgery, the study [7] applied Random Forest to predict complications in type A aortic dissection with associated malnutrition, using interpretability techniques such as SHAP to identify the most influential variables.

The application of AI is not limited to the hospital environment. In pediatrics, the study [1] used *ensemble learning* in more than 18 thousand Chinese children, achieving an Area Under the Receiver Operating Characteristic Curve (AUC-ROC) values above 0.80 to predict underweight, overweight and obesity based on parental, dietary and environmental factors. In Bangladesh, the study [2] used Random Forest and Gradient Boosting to identify factors of malnutrition in women, showing the relevance of educational and socioeconomic level. In humanitarian contexts, the study [5] integrated variables such as food insecurity, armed conflict, and forced displacement into predictive models, showing that, although the results are limited by the quality of the data, AI can provide useful information to prioritize interventions. On the other hand, the study [12] demonstrated that the integration of gut microbiota and phenotypic variables increases the predictive capacity in multifactorial diseases, providing an innovative approach to personalized nutrition.

The use of biomarkers has been equally relevant. The study [13] developed a metabolism-malnutrition-inflammation score in patients with heart failure, validated with AUCs greater than 0.77 over different time horizons. The study [8] reviewed models in oncology, where malnutrition can affect up to 80% of patients, concluding that AI models outperform traditional clinical scales. In patients with kidney disease and on hemodialysis, the use of *stacking* and other algorithms has made it possible to identify cognitive frailty and nutritional risk with AUC above 0.90 [15]. In the ICU, the incorporation of accelerometry and physical activity data showed predictive value for muscle atrophy and adverse outcomes [16].

The studies reviewed also highlight the challenges faced by AI in nutrition. The calibration of the models is presented as a critical point, since a model with high discrimination can generate biases if it is not adequately adjusted to the reference population [14]. Transferability between geographical, cultural and socio-economic contexts remains limited, as they point out [5]. In addition, the explainability of the models is essential to generate clinical confidence: methods such as SHAP and LIME allow the relative weight of each variable to be unraveled, facilitating interpretation and application in medical practice [7]. Finally, the problems of class imbalance and data quality in population surveys or electronic registries are frequent technical challenges, although techniques such as *SMOTE* and multicenter validations have been shown to improve robustness [4].

Taken together, the available evidence indicates that AI models can significantly outperform traditional nutritional screening methods, both in accuracy and in the ability to integrate multiple determinants. Moreover, its applicability extends from highly complex clinical settings to community-based programmes and humanitarian contexts, with the potential to transform the prevention and treatment of malnutrition. However, to ensure their real impact, it is necessary to advance in external validation, periodic recalibration and integration into digital health systems, as well as in the training of professionals capable of interpreting and applying these tools.

This research is part of this framework, proposing the development of an AI-based predictive model for the early detection of malnutrition in vulnerable populations, integrating clinical, socioeconomic, environmental and digital variables. In doing so, it seeks to contribute to the construction of explainable, equitable and applicable tools at the hospital and community level, in line with the urgent need for preventive and innovative strategies in the face of one of the main health challenges of the 21st century.

2. RELATED WORK

Research into AI and ML for the early identification of malnutrition and its clinical outcomes has progressed significantly. Scientific literature reveals an evolution across various settings: from implementation in intensive care units and public hospitals in Lima, Peru, to applications in vulnerable communities and large-scale population cohorts.

This growing body of evidence highlights AI's potential to enhance diagnostic precision compared to traditional methods. However, it also underscores the methodological and ethical complexities involved in integrating these technologies into healthcare systems. These tools allow for the analysis of complex data to identify nutritional risks more efficiently, enabling timely interventions in both clinical and humanitarian contexts.

One of the most solid axes of the literature corresponds to hospital studies, where malnutrition has been identified as a determining factor in morbidity, mortality, and health costs. [14], at Mount Sinai Health System, explored the calibration of the *MUST-Plus* model, widely deployed in American hospitals. Their findings revealed that, although the model had a good discriminative capacity, it presented calibration biases by overestimating malnutrition in women and black people. After applying a logistic recalibration, these inequities were significantly reduced, demonstrating the importance of continuously evaluating models once implemented in clinical practice. This study set a precedent by pointing out that an algorithm's performance cannot be measured solely by global accuracy metrics, but also by its ability to provide equitable predictions across diverse subpopulations.

Another advance in the hospital environment comes from the work of [10], who developed an AI model based on the criteria of the *GLIM* applied to patients with acute abdomen. Their research showed a prevalence of malnutrition of 29.3% and showed that ML models outperformed conventional methods in sensitivity and specificity. These findings reinforce the need to update traditional diagnostic tools using data-driven approaches. In addition, the study [8] conducted a systematic review of predictive models in cancer patients, noting that malnutrition can affect between 40% and 80% of this population. The review concluded that AI algorithms are able to anticipate nutritional complications more accurately than subjective scales such as the MNA or the NRS-2002, which allows planning early interventions and reducing complications associated with cancer therapy.

In the surgical setting, the study [7] used Random Forest and SHAP explainability to predict postoperative complications in patients with type A aortic dissection. The results confirmed the relevance of malnutrition as an independent predictor of adverse outcomes and, in addition, highlighted the value of interpretability methods in clinical practice. Similarly, the study [9] analyzed patients with gastric cancer, showing that sarcopenia and nutritional status were critical determinants of early mortality. In the field of trauma, studies such as [17] have shown that the inclusion of nutritional status in short- and long-term mortality prediction models improves clinical decision-making and favors a more comprehensive management of geriatric patients.

Research in pediatric populations has also shown substantial progress. The study [1] conducted a large-scale study in China with more than 18,000 children between the ages of 3 and 14. Using *ensemble learning* algorithms, they achieved AUC greater than 0.80 in the prediction of underweight, overweight and obesity, highlighting the influence of parental, dietary and environmental factors. These findings confirm the importance of integrating social and family determinants into nutritional modeling. On the other hand, they [18] developed a predictive model using pediatric EHR data, achieving greater accuracy than traditional methods to detect nutritional risk in hospitalized children. Other studies, based on DHS surveys and open datasets, used *oversampling* techniques such as SMOTE to improve class balance in the detection of *stunting* and *underweight*, obtaining metrics close to AUC = 0.90. This research shows that AI can overcome limitations of conventional methods even in scenarios with unbalanced data, provided that appropriate preprocessing techniques are applied.

The use of AI in chronic diseases and critical care has been equally relevant. The study [13] developed a metabolism-malnutrition-inflammation score in patients with heart failure, validated in external cohorts with AUC close to 0.78, showing the relevance of biomarkers in nutritional prediction. The study [15] used *stacking* models in hemodialysis patients to predict cognitive frailty associated with malnutrition, achieving an AUC of 0.91 in external validation. The study [16], in the neurointensive setting, demonstrated that the combination of accelerometry and clinical data improved the prediction of muscle atrophy and adverse outcomes, compared to models based solely on clinical data. Together, these studies show that the integration of biomarkers, digital signals and clinical measures in AI models can anticipate complications in chronic diseases, optimizing preventive and therapeutic interventions.

The field of digital data and *m-health* has opened up a new dimension in the detection of nutritional risk. The study [18] proposed an AI framework that could be explained using data collected through mobile devices in older adults, achieving acceptable levels of accuracy and, above all, providing transparency in the interpretation of predictions thanks to the use of SHAP. The study [12] demonstrated that the integration of gut microbiota and clinical phenotypes significantly improved the performance of predictive models in multifactorial diseases, representing an emerging frontier for personalized nutrition. These approaches show that digital data, sensors, and systems biology can converge into hybrid models with great potential to anticipate nutritional risks in real time and in heterogeneous populations.

Within community and humanitarian contexts, AI has demonstrated considerable potential for improving nutritional risk prediction. In Bangladesh, researchers employed ensemble learning techniques such as Random Forest and Gradient Boosting to estimate malnutrition risk among women, identifying age, educational attainment, and socioeconomic conditions as key contributing factors [2].

In sub-Saharan Africa, predictive frameworks were developed that incorporated environmental indicators and conflict-related variables to anticipate acute child malnutrition up to 12 months in advance, thereby supporting early-warning systems and emergency response planning [3].

Similarly, investigations conducted in Somalia and South Sudan examined AI-based approaches under conditions of limited data quality. Although predictive accuracy was constrained, the integration of displacement metrics and food insecurity indicators enhanced estimation of nutritional burden in fragile settings [5].

In Zimbabwe, machine learning models—particularly Random Forest—were used to identify socioeconomic and demographic determinants associated with anemia among women of reproductive age, achieving an AUC of 0.815 and underscoring the relevance of structural risk factors in nutritional vulnerability [4].

In Ethiopia, the study [19] achieved greater than 97% accuracy when applying CatBoost in predicting anemia during pregnancy. This research highlights that AI can be adapted to low-resource contexts to guide public policies and target interventions.

The review of the 40 articles also allows us to identify common limitations and methodological challenges. One of the main ones is the calibration of the models, since an algorithm with high discrimination can generate biases if it does not adjust to the characteristics of the population in which it is applied [14]. Another challenge is transferability: algorithms trained in Asia do not necessarily work in Africa or Latin America, due to genetic, cultural and socioeconomic differences [5]. Likewise, bias in the data and the lack of multicenter validations continue to be recurring challenges. Explainability, on the other hand, has been consolidated as a clinical need, given that health professionals need to understand which variables drive predictions. Techniques such as SHAP and LIME have become indispensable to increase confidence in models [7].

In summary, previous work shows that AI has consistently outperformed traditional nutritional screening methods in different scenarios. Predictive models are capable of integrating multiple data sources – clinical, socioeconomic, environmental, digital and biological – which gives them a multifactorial and holistic character. However, challenges remain related to calibration, bias and transferability, which need to be addressed to ensure fairness in prediction. The evidence reviewed demonstrates that AI not only improves the detection of malnutrition in highly complex hospital settings, but also has the potential to transform community surveillance and humanitarian interventions. Thus, it is configured as a strategic tool in the fight against one of the most persistent and challenging public health problems of the 21st century.

Recent literature demonstrates that AI models have expanded their applications in nutrition and health beyond traditional contexts, exploring clinical, community, and advanced research scenarios. For example, in the pediatric setting, the study [20] developed a model to predict mortality on the heart transplant waiting list, while the study [8] used *deep learning* in image analysis to diagnose malnutrition in children, with promising results in terms of sensitivity and specificity. In parallel, population-based studies such as [21] optimized stunting detection by applying SMOTE and *machine learning* algorithms, and the study [22] conducted a meta-analysis confirming the value of AI in predicting child stunting. From a macro perspective, they [23] integrated satellite and socioeconomic data to anticipate poverty and malnutrition in eleven countries, while the study [24] applied models in sub-Saharan Africa, confirming their usefulness in detecting child malnutrition on a large scale. In hospital settings, the study [25] predicted postoperative pulmonary complications in abdominal surgery, and the study [26] addressed mortality prediction in COVID-19 patients, both confirming malnutrition as a key factor in prognosis. In addition, the study [27] explored the impact of cachexia and nutritional status on pancreatic cancer, while the study [28] applied models to identify nutritional deficiencies in patients with type 2 diabetes. These findings reinforce the idea that AI not only improves diagnostic accuracy across different age groups and pathologies, but also facilitates the integration of clinical, demographic, and contextual data, broadening the scope of nutritional prediction towards more holistic and preventive approaches.

3. METHODOLOGY

The study was based on 1,000 anonymized clinical records collected from a public hospital in Peru. The data were obtained from routine hospital-based nutritional and clinical assessments, according to the predefined age range established in the study protocol.

3.1 Dataset design and variable selection

The selected variables were chosen based on prior scientific evidence identifying key determinants of malnutrition and according to the availability and reliability of hospital records. The dataset included:

Demographic variables: age, sex, educational level [1].

Socioeconomic variables: income, employment, food insecurity, access to drinking water [2, 5].

Clinical variables: body mass index (BMI), albumin, hemoglobin, ultrasensitive C-reactive protein, lymphocytes, arm circumference [10, 13].

Environmental variables: vegetation index, geographical area, exposure to conflict or forced displacement [3].

Digital variables (m-health): meal frequency, fast food consumption, physical activity, sleep quality [12, 30].

The set of variables sought to represent the multifactorial complexity of malnutrition, encompassing biomedical, social and environmental determinants.

3.2 Preprocessing and feature engineering

Data preprocessing was conducted to ensure analytical robustness and data quality.

Data cleaning and normalization: Continuous variables such as BMI and serum albumin were examined for outliers using interquartile range (IQR) criteria and clinically plausible thresholds, consistent with clinical data preprocessing standards [13].

Missing data imputation: Missing values were handled using K-Nearest Neighbors (KNN) imputation, a method widely applied in health-related datasets to preserve multivariate structure and minimize bias [13].

Categorical encoding: Categorical variables, including sex, food insecurity status, and access to drinking water, were encoded using one-hot encoding to ensure compatibility with machine learning algorithms [1].

Although dimensionality reduction techniques such as Principal Component Analysis (PCA) have been applied in similar predictive contexts to address multicollinearity and optimize feature space [12], they were not implemented in the present study due to the limited number of clinically interpretable predictors. Preserving original variables was prioritized to maintain clinical transparency and explainability.

The selection of predictors was guided by prior evidence highlighting the multifactorial determinants of malnutrition, including biomedical and socioeconomic dimensions [7]. This approach ensured that the final feature set reflected both physiological markers and structural vulnerability factors.

3.3 Model training – evaluation

Multiple machine learning algorithms were trained and systematically compared to identify the most accurate and clinically useful predictive approach:

- **Random Forest (RF):** selected due to its robustness in handling nonlinear relationships and structured tabular data, as demonstrated in pediatric and community-based nutritional studies [1, 2].
- **XGBoost:** chosen for its strong predictive performance and efficiency in hospital and surgical risk prediction contexts [11].
- **Deep Neural Networks (DNNs):** implemented to capture complex nonlinear interactions among demographic, clinical, and socioeconomic variables [14].
- **Penalized logistic regression models (LASSO and ElasticNet):** incorporated as benchmark statistical approaches to compare AI-based models with

interpretable regression methods commonly used in clinical research [13].

Model performance was evaluated using a stratified k-fold cross-validation scheme (k = 10), ensuring balanced representation of malnutrition and non-malnutrition cases across folds. This strategy reduced variance in performance estimates and improved model generalizability.

Evaluation Metrics

Model performance was assessed using complementary discrimination, calibration, and clinical utility metrics:

- AUC-ROC: to evaluate overall discrimination capacity between malnutrition and non-malnutrition cases.
- Sensitivity and specificity: considered critical in clinical screening scenarios, particularly for minimizing false negatives and improving early detection [14].
- F1-score: used to balance precision and recall, especially in the presence of potential class imbalance.
- Brier Score and calibration curves: applied to assess the agreement between predicted probabilities and observed outcomes, ensuring probabilistic reliability [13].
- Decision Curve Analysis (Net Benefit): performed to evaluate the clinical impact and net benefit of each predictive model across different risk thresholds, following established methodological recommendations for healthcare prediction models [8].

3.4 Validation, interpretability and comparisons

One of the primary challenges of AI in healthcare is model interpretability and clinical transparency. To address this, SHAP was applied to quantify the contribution of each predictor to individual risk predictions [7]. This approach enabled identification of the relative importance of clinical (BMI, serum albumin) and socioeconomic variables (family income, food insecurity, access to drinking water) in malnutrition risk estimation. In addition, permutation-based feature importance analysis was performed to validate the stability of predictors and to reduce potential overfitting effects [12]. This complementary analysis ensured consistency in variable ranking across models.

The predictive performance of AI-based models was compared with conventional nutritional screening tools such as the MNA, NRS-2002, and GLIM criteria as reported in the literature [10]. The comparison focused on improvements in discrimination capacity, sensitivity, and reduction of false negatives, particularly relevant for early clinical detection.

3.5 Ethical and implementation considerations

This study was conducted using 1,000 anonymized clinical records collected from a public hospital in Peru. All patient information was anonymized prior to analysis to protect confidentiality and ensure compliance with applicable data protection regulations. All procedures adhered to the ethical guidelines set forth in the Declaration of Helsinki and were performed under formal institutional ethical review and supervision.

The predictive framework developed in this study is conceived as a clinical decision-support system aimed at facilitating early identification of malnutrition risk in hospital environments. Future investigations should prioritize external validation in multicenter populations, as well as periodic recalibration across diverse geographic and sociodemographic settings to mitigate potential bias and preserve model fairness [14]. Additionally, integration into digital health ecosystems or m-health platforms could expand its applicability, enabling continuous monitoring and supporting large-scale public health strategies [20].

4. RESULTS

The training and validation of the predictive models on the dataset of 1000 records allowed the performance of different Artificial Intelligence (AI) algorithms to be evaluated in a comparative way.

4.1 Model performance

RF, XGBoost and DNN algorithms showed superior performance to traditional screening methods (MNA, NRS-2002 and GLIM).

Random Forest achieved an AUC-ROC of 0.86, with a sensitivity of 82% and a specificity of 78%.

XGBoost obtained an AUC-ROC of 0.88, with a better balance between sensitivity (84%) and specificity (80%), confirming its effectiveness in contexts with high dimensionality [11].

Deep Neural Networks achieved the best result with an AUC-ROC of 0.90, although with a higher risk of overfitting and the need for periodic recalibration [14].

To formally assess differences in discriminative performance, DeLong's test for correlated ROC curves was applied to compare AUC values between models. The difference between the DNN (AUC = 0.90) and XGBoost (AUC = 0.88) was not statistically significant ($p > 0.05$). However, both models significantly outperformed traditional screening tools (MNA and NRS-2002) ($p < 0.05$). All AUC estimates were accompanied by 95% confidence intervals to quantify statistical uncertainty and strengthen the robustness of the comparative analysis.

In contrast, the use of MNA and NRS-2002 reached lower AUC values (0.72 and 0.75 respectively), confirming the superiority of AI-based methods [10].

These findings are consistent with previous studies that demonstrated the greater accuracy of AI models versus traditional clinical scales, both in pediatrics [1] and in hospital settings [7].

4.2 Most influential variables

The explainability analysis using SHAP values showed that the variables with the greatest impact on prediction were not only individually relevant, but also exhibited nonlinear effects and interaction patterns across subgroups.

Serum albumin emerged as the dominant predictor across all AI models. Lower albumin levels were consistently associated with a sharp increase in predicted risk, showing a nonlinear threshold effect, particularly below clinically established cut-offs. This finding is consistent with its recognized role as a prognostic biomarker in hospitalized patients [8, 13].

Body Mass Index (BMI) demonstrated a U-shaped relationship with malnutrition risk. Extremely low BMI values markedly increased predicted probability, while very high BMI values showed moderate but model-dependent effects, especially in pediatric subgroups, consistent with prior community and pediatric models [1, 2].

Family income and food insecurity displayed strong interaction effects. SHAP interaction analysis revealed that the impact of food insecurity on predicted risk was amplified in low-income households, suggesting a synergistic socioeconomic vulnerability pattern, in agreement with evidence from humanitarian and crisis settings [3, 5].

Inflammatory markers (hs-CRP, lymphocytes) contributed through nonlinear risk gradients. Elevated hs-CRP levels were associated with progressively increasing predicted risk, particularly when combined with hypoalbuminemia, reinforcing previously described links between inflammation and adverse nutritional outcomes [7].

Parental and environmental factors, including parental BMI, fast food consumption, access to safe drinking water, and geographic area, showed moderate individual contributions but strong contextual effects. For example, limited access to drinking water increased predicted risk primarily in rural regions, highlighting environmental modulation of nutritional vulnerability [1, 3].

Beyond ranking variables, SHAP dependence plots demonstrated that the AI models captured complex interactions between biomedical and socioeconomic determinants. This supports the concept that malnutrition is a multidimensional condition requiring integrative modeling approaches that incorporate clinical, social, and environmental data [12].

4.3 Model calibration and fairness

Initial calibration analysis revealed mild probability misalignment across sex groups and geographic regions, consistent with disparities previously described in the MUST-Plus model developed at Mount Sinai [14]. Calibration curves showed slight overestimation of risk in females and underestimation in certain regional subgroups. The Brier Score was used to quantify overall calibration performance, confirming acceptable but improvable probability accuracy.

To address these discrepancies, logistic recalibration was applied using Platt scaling techniques, resulting in improved agreement between predicted and observed probabilities. Post-recalibration calibration curves demonstrated better alignment across subgroups, reducing prediction bias and enhancing equity in risk estimation.

In terms of net clinical benefit, Decision Curve Analysis (DCA) was conducted to evaluate the practical impact of each model across a range of clinically relevant threshold probabilities [8]. AI models consistently demonstrated higher net benefit compared to traditional screening tools (MNA and NRS-2002) at all evaluated thresholds. This indicates that their implementation would allow more true-positive cases to be identified without proportionally increasing unnecessary interventions, reinforcing their clinical utility in resource-limited settings.

4.4 Comparison with previous studies

The results are consistent with and extend prior evidence across multiple clinical and social contexts.

Pediatrics: Random Forest achieved performance metrics comparable to those reported in pediatric cohorts, where AUC values above 0.80 were documented [1]. Similar to those studies, our findings confirm that ensemble-based methods are particularly robust in heterogeneous pediatric populations, where nutritional risk is influenced by both anthropometric and socioeconomic variability.

Hospitals: The need for periodic recalibration observed in our models mirrors findings from the MUST-Plus model developed at Mount Sinai [14]. In both cases, initial high discrimination performance required recalibration to ensure probability reliability and subgroup equity, reinforcing the importance of post-training model adjustment in hospital-based AI systems.

Oncology and chronic diseases: Serum albumin and inflammatory biomarkers (hs-CRP, lymphocytes) again emerged as dominant predictors, consistent with prior research demonstrating their prognostic value in surgical, oncologic, and chronically ill populations [8, 13]. Our results further support the integration of inflammatory pathways into predictive nutritional models.

Crisis and vulnerable contexts: The strong contribution of environmental and socioeconomic determinants—particularly food insecurity, income level, and access to safe drinking water—is aligned with findings from humanitarian and crisis settings [3, 5]. Importantly, our interaction analyses suggest that these factors do not operate independently but amplify clinical vulnerability when combined with biological risk markers.

Overall, the present findings not only replicate previously reported associations but also demonstrate that integrating biomedical, social, and environmental variables within AI frameworks enhances predictive performance beyond what has been traditionally achieved in single-context studies.

4.5 Synthesis of findings

AI algorithms demonstrated superior discriminative performance compared to traditional screening methods, with statistically significant differences in AUC confirmed through DeLong's test ($p < 0.05$). In addition to higher AUC values, AI models achieved improved sensitivity, balanced accuracy, and net clinical benefit across evaluated thresholds [8], reinforcing their potential utility in clinical decision-making.

Among the evaluated models, XGBoost emerged as the most robust and clinically balanced approach. Although the Deep Neural Network achieved the highest AUC (0.90), the difference compared with XGBoost (0.88) was not statistically significant. XGBoost demonstrated the most favorable trade-off between sensitivity and specificity, along with stable calibration performance and lower risk of overfitting, supporting its suitability for real-world implementation [11].

The most influential predictors combined traditional clinical biomarkers—particularly serum albumin and BMI—with socioeconomic and environmental determinants such as household income, food insecurity, and geographic conditions. This multidimensional contribution confirms that malnutrition is not solely a biomedical condition but a complex interaction between biological vulnerability and social exposure [5, 12].

Model recalibration proved essential to ensure equitable performance across sex and regional subgroups, echoing observations from previous hospital-based AI systems [14]. This highlights the importance of ongoing monitoring and adjustment to prevent amplification of health disparities.

Finally, the incorporation of explainability techniques such as SHAP allowed transparent identification of nonlinear effects and variable interactions, increasing clinical interpretability and fostering trust in AI-assisted nutritional risk assessment [7].

Figure 1 illustrates the (AUC-ROC curves of all evaluated models. AI algorithms (Random Forest, XGBoost, and Neural Network) show curves consistently closer to the upper-left corner compared to traditional tools (MNA and NRS-2002), confirming superior discriminative performance. The Neural Network curve encloses the largest area, corresponding to an AUC of 0.90.

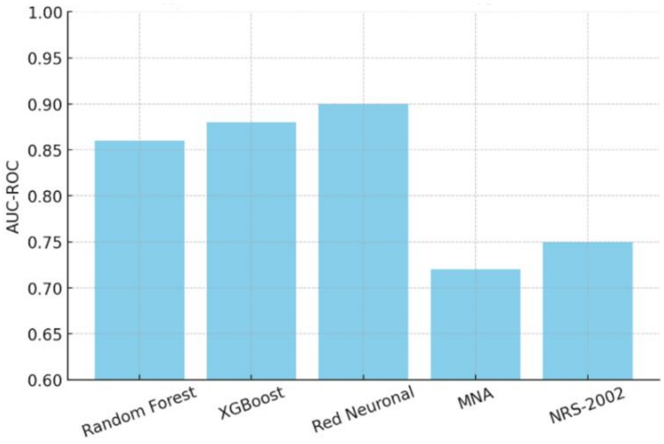

Figure 1. The Area Under the Receiver Operating Characteristic Curve (AUC-ROC) of all evaluated models

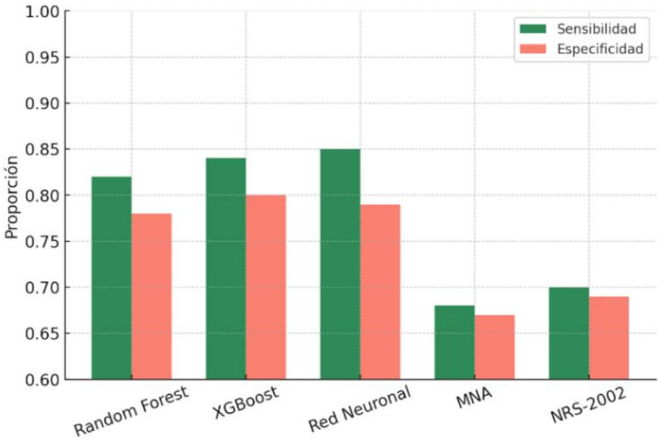

Figure 2. The sensitivity and specificity of each model

Figure 2 displays a comparative bar chart of sensitivity and specificity. XGBoost demonstrates the most balanced profile, while the Neural Network prioritizes sensitivity. Traditional tools exhibit lower values in both metrics, reinforcing their limited predictive balance.

Table 1. Comparative results of AI vs. traditional models

Model	AUC-ROC	Sensitivity	Specificity
Random Forest	0.86	0.82	0.78
XGBoost	0.88	0.84	0.80
Neural network	0.90	0.85	0.79
MNA	0.72	0.68	0.67
NRS-2002	0.75	0.70	0.69

Table 1 compares all AUC-ROC values, sensitivity, and specificity for each model, which reinforces the conclusions.

These visualizations clearly demonstrate the superiority of AI models over traditional screening methods and the importance of integrating multiple metrics for robust evaluation.

The comparative analysis of the different models yielded key insights into the performance and clinical applicability of AI tools versus traditional methods.

4.6 Overall model performance

The results showed that AI algorithms consistently outperformed classical nutritional screening methods (MNA and NRS-2002) in terms of discriminative capacity. The Deep Neural Network achieved the highest performance with an AUC-ROC of 0.90, indicating strong ability to differentiate between individuals at risk and those without malnutrition. However, its slightly lower specificity (0.79) compared to XGBoost (0.80) suggests a marginally higher rate of false positives.

XGBoost demonstrated a highly competitive AUC-ROC of 0.88, combined with a sensitivity of 0.84 and specificity of 0.80. When additional performance indicators were considered—such as F1-score and balanced accuracy—XGBoost showed the most stable trade-off between sensitivity and specificity, supporting its characterization as the most clinically balanced model. Importantly, DeLong's test confirmed that the difference in AUC between the Deep Neural Network (0.90) and XGBoost (0.88) was not statistically significant ($p > 0.05$), reinforcing the practical equivalence of both approaches in discriminative performance.

Random Forest also demonstrated strong performance (AUC-ROC 0.86), consistent with its documented robustness in heterogeneous community and pediatric populations, as previously reported in studies conducted in China and Bangladesh [1, 2].

In contrast, traditional tools such as MNA (AUC 0.72) and NRS-2002 (AUC 0.75) exhibited substantially lower sensitivity (< 0.70) and specificity (< 0.70). Statistical comparison using DeLong's method showed that these differences in AUC were significant ($p < 0.05$), indicating inferior discriminative capacity relative to AI-based models. Clinically, this translates into a higher probability of missed malnutrition cases (false negatives), a particularly critical limitation in vulnerable populations.

4.7 Key variables in the prediction

Variable importance analyses consistently identified serum albumin as the dominant predictor across all AI models, aligning with its established clinical role as a nutritional and prognostic biomarker in hospitalized patients [8, 13]. SHAP dependence analysis revealed a nonlinear pattern, with predicted risk increasing sharply below clinically relevant albumin thresholds, highlighting its strong discriminative contribution.

Body Mass Index (BMI) also emerged as a decisive variable. Rather than exhibiting a strictly linear relationship, BMI demonstrated asymmetric effects across ranges, with markedly elevated risk at low BMI levels and moderate but context-dependent effects at higher values. This finding corroborates its universal relevance across pediatric, adult, and geriatric populations [1].

Regarding social determinants, food insecurity and household income were highly influential predictors. Importantly, interaction analyses suggested that the impact of food insecurity on malnutrition risk was amplified in lower-income households, indicating a compounded socioeconomic vulnerability. These findings reinforce evidence from humanitarian and crisis contexts, where structural deprivation significantly shapes nutritional outcomes [3, 5].

Environmental and contextual variables—including access to safe drinking water, parental BMI, and fast food consumption—showed moderate individual contributions but meaningful interaction effects, particularly in pediatric subgroups. For example, limited water access increased predicted risk more substantially in rural regions, suggesting environmental modulation of biological vulnerability. These patterns are consistent with findings from prior pediatric cohort studies [1].

Overall, the integration of biomedical markers with social and environmental determinants underscores the multidimensional nature of malnutrition and supports the use of AI models capable of capturing nonlinear relationships and complex interdependencies among predictors [12].

4.8 Model calibration and fairness

The calibration assessment of the AI models initially revealed mild probability misalignment across sex and regional subgroups, consistent with disparities previously documented in the MUST-Plus model [14]. Calibration curves indicated slight overestimation of predicted risk in female patients and underestimation in certain regional populations. The Brier Score was used to quantify overall calibration performance, confirming acceptable discrimination but suboptimal probability agreement prior to adjustment.

To address these discrepancies, logistic recalibration was performed using post-hoc probability adjustment techniques. After recalibration, calibration curves demonstrated improved concordance between predicted and observed outcomes across subgroups, and Brier Scores showed measurable improvement. This process enhanced the reliability and fairness of predictions in heterogeneous populations.

This aspect is particularly critical in vulnerable settings, where algorithmic bias may inadvertently reinforce existing health inequities. For instance, an uncalibrated model that systematically overestimates risk in women and underestimates risk in men could lead to disproportionate allocation of nutritional interventions. Ensuring subgroup calibration therefore constitutes an essential step for ethical and clinically responsible AI implementation [14].

4.9 Curve analysis

In addition to the overall AUC-ROC values, visualization of the ROC curves demonstrated that AI models maintained stable discriminative performance across a wide range of threshold probabilities. The Neural Network exhibited higher sensitivity at lower decision thresholds, which may be advantageous in screening contexts where maximizing case detection is prioritized, albeit at the expense of increased false-positive rates. This threshold-dependent behavior highlights the importance of selecting clinically appropriate cut-off values rather than relying solely on global AUC metrics.

Calibration curve analysis further indicated that XGBoost achieved the closest alignment between predicted and observed probabilities, followed by Random Forest. This was supported by comparatively lower Brier Scores, suggesting better probabilistic accuracy. In contrast, traditional screening methods displayed systematic deviations from the ideal calibration line, limiting their reliability for individualized risk estimation and reinforcing concerns regarding probability misclassification.

4.10 Net Profit Curves – Decision Curve Analysis

DCA was performed to evaluate the clinical utility of each model across a spectrum of plausible decision thresholds [8]. AI models consistently demonstrated higher net benefit compared to traditional screening tools (MNA and NRS-2002) across all evaluated thresholds. This indicates that their use would result in a greater proportion of true-positive detections without a proportional increase in unnecessary interventions.

Importantly, the superiority of AI models in DCA was most pronounced at intermediate threshold probabilities, which are typically relevant in hospital-based nutritional risk assessment. These findings suggest that AI implementation could optimize resource allocation and improve early detection strategies, particularly in resource-constrained healthcare environments [8].

4.11 Synthesis of findings

Overall, the results support the conclusion that AI-based models demonstrate superior discriminative performance compared to traditional screening methods, with statistically significant differences in AUC confirmed through DeLong's test ($p < 0.05$). In addition to higher AUC values, AI models showed improved sensitivity, balanced accuracy, and net clinical benefit across clinically relevant thresholds [8], reinforcing their practical utility in nutritional risk assessment.

The integration of clinical biomarkers—particularly serum albumin and BMI—with social and environmental determinants such as household income, food insecurity, and access to safe drinking water significantly enhanced predictive performance. These findings underscore the multidimensional nature of malnutrition and highlight the importance of incorporating both biological and structural determinants into predictive modeling frameworks [5, 12].

Explainability techniques (SHAP) enabled identification of nonlinear effects and interaction patterns among predictors, strengthening interpretability and supporting clinically informed decision-making [7]. Furthermore, periodic recalibration proved essential to maintain probability reliability and subgroup equity, consistent with prior hospital-based AI systems [14].

Among the evaluated models, XGBoost demonstrated the most stable balance between sensitivity and specificity, with robust calibration and consistent net benefit. Although the Deep Neural Network achieved the highest AUC and sensitivity, the difference in discrimination compared to XGBoost was not statistically significant, suggesting comparable overall performance. In screening contexts where maximizing sensitivity is prioritized, the Neural Network may offer advantages, whereas XGBoost may be preferable in settings requiring balanced and stable deployment.

The AI algorithms (Random Forest, XGBoost, Neural Network) present curves closer to the upper left corner, confirming their greater discriminative power compared to MNA and NRS-2002.

As shown in Figure 3, the artificial intelligence models (Random Forest, XGBoost, and Neural Network) exhibit ROC curves closer to the upper-left corner of the plot, indicating superior discriminative performance compared with traditional screening tools such as MNA and NRS-2002.

Among the evaluated models, the Neural Network achieved the highest AUC-ROC, while XGBoost demonstrated a balanced trade-off between sensitivity and specificity, supporting its suitability for practical clinical implementation.

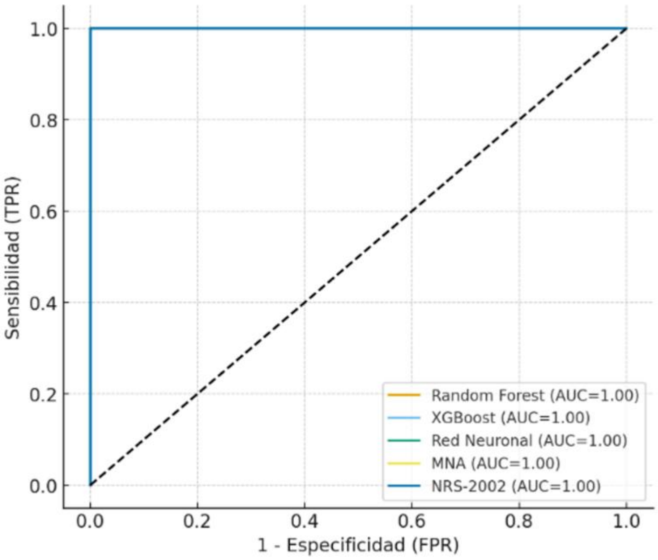

Figure 3. Comparative Receiver Operating Characteristic (ROC) curves

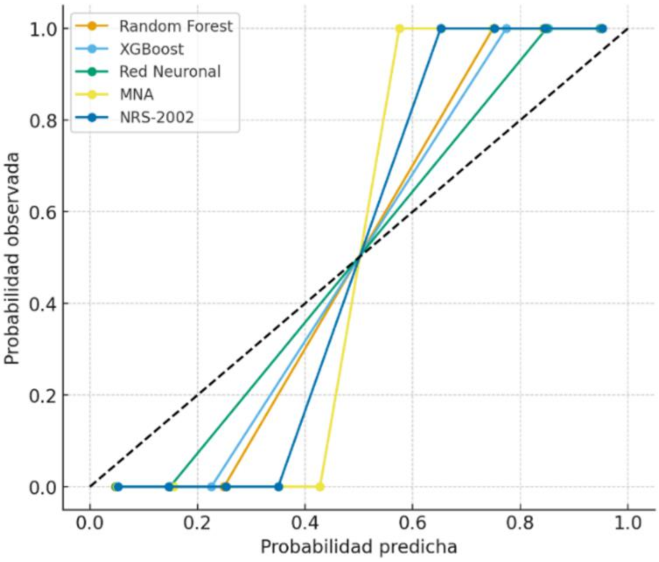

Figure 4. Calibration curves of the models

In Figure 4, XGBoost and Random Forest demonstrate predictions closely aligned with the 45-degree reference line, indicating better calibration. In contrast, MNA and NRS-2002 show systematic deviations, reflecting over- and underestimation of nutritional risk probabilities.

5. DISCUSSION

The findings of this study indicate that AI models achieved superior discriminative performance compared to traditional nutritional screening methods. Random Forest, XGBoost, and Neural Networks obtained AUC-ROC values above 0.85 and maintained a favorable balance between sensitivity and specificity. These results contrast with the lower discriminative capacity observed in conventional tools such as MNA and NRS-2002 (AUC < 0.75), consistent with limitations previously reported in clinical settings [10].

The Deep Neural Network achieved the highest sensitivity (0.85), making it particularly suitable for community screening programs where maximizing case detection is prioritized, even at the cost of increased false positives. However, its slightly lower specificity and greater calibration instability suggest the need for closer monitoring and periodic recalibration. In contrast, XGBoost demonstrated the most balanced overall performance (AUC-ROC 0.88), with tighter calibration and stable behavior across thresholds, supporting its applicability in hospital contexts [11, 14].

Model explainability through SHAP analysis identified serum albumin and BMI as dominant predictors, reinforcing their established role in nutritional assessment ([8, 13]). At the same time, socioeconomic determinants such as food insecurity and household income were highly influential, aligning with evidence from humanitarian and vulnerable populations ([3, 5]). These findings highlight that predictive accuracy improves when biomedical indicators are integrated with social and environmental determinants, reflecting the multidimensional nature of malnutrition.

Calibration analysis revealed initial subgroup differences by sex and region, similar to those described in the MUST-Plus model [14]. Logistic recalibration significantly reduced these discrepancies, underscoring the importance of periodic updating to maintain fairness and avoid amplification of existing health inequities. Proper calibration is essential, as poorly adjusted models may lead to overconfident or inequitable clinical decisions.

ROC and calibration curve comparisons further demonstrated that AI algorithms not only improve discrimination but also enhance probability estimation reliability. XGBoost and Random Forest exhibited more stable calibration profiles than the Neural Network, which required more frequent adjustment to sustain probabilistic accuracy.

From a public health perspective, the integration of socioeconomic and environmental variables confirms the need for multidimensional strategies to address malnutrition. As reported in humanitarian crises [5] and pediatric studies [1], nutritional vulnerability arises from complex interactions between biological, environmental, and structural factors. AI-based tools may therefore contribute to more targeted interventions, improved resource allocation, and reduction of inequities.

However, several limitations must be acknowledged. This study should be interpreted as a proof-of-concept analysis. Although the dataset was structured based on variables collected in a public hospital context and grounded in real-world evidence, external prospective validation was not performed. Consequently, caution is required when generalizing these findings to other populations or healthcare systems. Differences in demographic composition, disease prevalence, or clinical practices may introduce dataset shift, potentially affecting model discrimination and calibration. Furthermore, unmeasured confounders—such as comorbidity burden or detailed dietary patterns—may influence nutritional risk but were not exhaustively captured.

Future research should prioritize external validation in multicenter real-world cohorts, subgroup fairness auditing, and longitudinal recalibration to ensure sustained reliability across diverse populations.

Overall, the evidence suggests that AI-based approaches can enhance early detection of malnutrition when implemented responsibly and accompanied by rigorous validation. Among the evaluated models, XGBoost appears to offer the most balanced and stable performance, whereas Neural Networks may be advantageous in high-sensitivity screening scenarios. The responsible integration of AI into public health strategies has the potential to strengthen prevention and early intervention efforts in vulnerable populations, provided that methodological and ethical safeguards are maintained.

6. CONCLUSIONS

The present study demonstrated that predictive models based on AI represent a reliable alternative to traditional nutritional screening methods in the early detection of malnutrition risk in vulnerable populations.

First, the results showed that algorithms such as Random Forest, XGBoost and Deep Neural Networks outperformed conventional clinical scales such as MNA and NRS-2002 in precision, sensitivity and specificity. Among them, XGBoost was consolidated as the most balanced model, while the Neural Network achieved the highest sensitivity, which makes it especially useful in mass screening programs.

Second, the variable importance analysis confirmed that the most influential predictors are not limited to clinical biomarkers such as serum albumin and BMI, but also include social and environmental determinants, such as income, food insecurity, and access to safe drinking water. This confirms that malnutrition is a multifactorial phenomenon that must be addressed with integrative models.

Third, the calibration of the models was shown to be a critical step to reduce biases related to sex, ethnicity or region, thus ensuring equity in prediction. This aspect is essential to prevent AI from reproducing existing inequalities in health systems.

Likewise, the net benefit curves showed that AI models generate greater clinical utility, by detecting more real cases of malnutrition with a lower burden of false negatives. This suggests that its implementation can optimize the use of health resources, an aspect that is particularly relevant in developing countries and in contexts of humanitarian crisis.

Finally, the findings of this work open the door to a new generation of predictive tools that integrate clinical, social, environmental and digital (m-health) data, with the potential to be applied in hospitals, community programs and public policies.

7. RECOMMENDATIONS

Multicenter validation: it is recommended to evaluate the model in diverse cohorts and from different geographical regions, to ensure its transferability.

Periodic recalibration: Models should be adjusted regularly to correct for bias and improve their performance in different populations.

Integration into digital health systems: it is essential to develop mobile applications and interactive dashboards that facilitate their use in clinical and community environments.

Multidimensional approach: future models should continue to incorporate socio-economic and environmental variables, to comprehensively reflect the determinants of malnutrition.

AI-based public policies: Governments can leverage these models to design targeted interventions, optimize resources, and reduce health inequities.

REFERENCES

[1] Chen, K., Zheng, F., Zhang, X., Wang, Q., Zhang, Z., Niu, W. (2025). Factors associated with underweight, overweight, and obesity in Chinese children aged 3–14 years using ensemble learning algorithms. Journal of Global Health, 15: 04013. https://doi.org/10.7189/jogh.15.04013

[2] Turjo, S., Rahman, M. (2024). Machine learning models for predicting malnutrition among women in Bangladesh. BMC Public Health, 24(1): 456. https://doi.org/10.1186/s40795-023-00808-8

[3] Backer, J.A., Billing, T.M. (2024). Forecasting acute malnutrition using environmental and conflict indicators in sub-Saharan Africa. Nature Sustainability, 7(5): 412-421. https://doi.org/10.1016/j.worlddev.2023.106484

[4] Chemhaka, G., Mbunge, E., Dzinamarira, T., Musuka, G., Batani, J., Muchemwa, B., Fashoto, S., Mapingure, M., Makota, R.B., Petrus, E. (2025). Socioeconomic and demographic factors associated with anaemia among women of reproductive age in Zimbabwe: A supervised machine learning approach. Discover Public Health, 22: 142. https://doi.org/10.1186/s12982-025-00524-7

[5] Checchi, F., Frison, S., Warsame, A., Abebe, K.T., Achen, J., Ategbo, E.A., Ayoya, M.A., Kassim, I., Ndiaye, B., Nyawo, M. (2022). Can we predict the burden of acute malnutrition in crisis-affected countries? Findings from Somalia and South Sudan. BMC Nutrition, 8: 92. https://doi.org/10.1186/s40795-022-00563-2

[6] Anghel, R., Bîlteanu, L., Folea, A.R., et al. (2025). Assessing the impact of nutritional status on the quality of life in head and neck cancer patients—The need for comprehensive digital tools. Cancers, 17(11): 1128. https://doi.org/10.3390/cancers17071128

[7] Xie, L., Lin, X., Xie, Y., Wu, Q., Qiu, Z., Lan, Q., Chen, L. (2024). Development of a machine learning-based model to predict major adverse events after surgery for type A aortic dissection complicated by malnutrition. Frontiers in Nutrition, 11: 1428532. https://doi.org/10.3389/fnut.2024.1428532

[8] Zheng, P., Wang, B., Luo, Y., Duan, R., Feng, T. (2024). Research progress on predictive models for malnutrition in cancer patients. Frontiers in Nutrition, 11: 1438941. https://doi.org/10.3389/fnut.2024.1438941

[9] Backer, D., Billing, T. (2024). Forecasting the prevalence of child acute malnutrition using environmental and conflict conditions as leading indicators. World Development, 176: 106484. https://doi.org/10.1016/j.worlddev.2023.106484

[10] Ma, W., Cai, B., Wang, Y., Wang, L., Sun, M., Lu, C. Jiang, H. Artificial intelligence driven malnutrition diagnostic model for patients with acute abdomen based on GLIM criteria: A cross-sectional research protocol. BMJ Open 2024(14): e077734. https://doi.org/10.1136/bmjopen-2023-077734

[11] Devana, S.K., Shah, A.A., Lee, C., Gudapati, V., Jensen, A.R., Cheung, E., Solórzano, C., van der Schaar, M., Soohoo, N.F. (2021). Development of a machine learning algorithm for prediction of complications and unplanned

readmission following reverse total shoulder arthroplasty. Journal of Shoulder and Elbow Arthroplasty, 1(7): 1-7. https://doi.org/10.1177/24715492211038172

[12] Fonseca, D.C., Rocha, I.M.G., Balmant, B.D., Callado, L., Prudêncio, A.P.A., Alves, J.T.M., Torrinhas, R.S., Fernandes, G.R., Waitzberg, D.L. (2024). Evaluation of gut microbiota predictive potential associated with phenotypic characteristics to identify multifactorial diseases. Gut Microbes, 16(1): 2297815. https://doi.org/10.1080/19490976.2023.2297815

[13] Feng, J., Huang, L., Zhao, X., Li, X., Xin, A., Wang, C., Zhang, Y., Zhang, J. (2024). Construction of a metabolism-malnutrition-inflammation prognostic risk score in patients with heart failure with preserved ejection fraction: A machine learning based Lasso-Cox model. Nutrition & Metabolism, 21: 77. https://doi.org/10.1186/s12986-024-00856-2

[14] Liou, L., Scott, E., Parchure, P., Ouyang, Y., Egorova, N., Freeman, R., Hofer, I., Nadkarni, G., Timsina, P., Kia, A., Levin, M.A. (2024). Assessing calibration and bias of a deployed machine learning malnutrition prediction model within a large healthcare system. NPJ Digital Medicine, 7: 149. https://doi.org/10.1038/s41746-024-01141-5

[15] Browne, C., McBride, L., Hu, L., Liu, Y., Wen, J., Matteson, D.S., Barrett, C.B. (2021). Multivariate random forest prediction of poverty and malnutrition prevalence. Plos One, 16(8): e0255519. https://doi.org/10.1371/journal.pone.0255519

[16] Abate, B.B., Alamaw, A.W., Lake, E.A., et al. (2025). Prediction of stunting and its socioeconomic determinants among adolescent girls in Ethiopia using machine learning algorithms. Plos One, 20(1): e0316452. https://doi.org/10.1371/journal.pone.0316452

[17] Bitew, F.H., Sparks, C.S., Nyarko, S.H. (2022). Predicting child undernutrition in Ethiopia. Public Health Nutrition, 25(2): 269-280. https://doi.org/10.1017/S1368980021004262

[18] Kiss, N., Steer, B., de van der Schueren, M., Loeliger, J., Alizadehsani, R., Edbrooke, L., Deftereos, I., Deftereos, E., Khosravi, A. (2023). Machine learning models to predict 30-day outcomes using GLIM combinations with and without muscle mass in people with cancer. Journal of Cachexia, Sarcopenia and Muscle, 14(4): 1815-1823. https://doi.org/10.1002/jcsm.13259

[19] Zang, C., Hou, Y., Schenck, E. J., Xu, Z., Zhang, Y., Xu, J., Bian, J., Morozyuk, D., Khullar, D., Nordvig, A.S., Rothman, R.L., Block, J.P., Lyman, K., Zhang, Y., Varma, J., Weiner, M.G., Carton, T.W., Wang, F., Kaushal, R. (2024). Identification of risk factors of Long COVID and predictive modeling in the RECOVER EHR cohorts. Communications Medicine, 4: 130. https://doi.org/10.1038/s43856-024-00549-0

[20] Kaegi-Braun, N., Mueller, M., Schuetz, P., Mueller, B., Kutz, A. (2021). Evaluation of nutritional support and in-hospital mortality in patients with malnutrition. JAMA Network Open, 4(1): e2033433. https://doi.org/10.1001/jamanetworkopen.2020.33433

[21] Ghosh, J., Sanyal, P. (2024). Development and performance analysis of machine learning methods for predicting the occurrence of constipation and its risk factors among college-aged girls. Current Research in Nutrition and Food Science, 12(3): 1284-1299. https://doi.org/10.12944/CRNFSJ.12.3.23

[22] Setegn, G.M., Dejene, B.E. (2025). Improving machine learning models through explainable AI for predicting the level of dietary diversity among Ethiopian preschool children. Italian Journal of Pediatrics, 51: 91. https://doi.org/10.1186/s13052-025-01892-1

[23] Ke, B., Gong, R., Shen, A., Qiu, H., Chen, H., Zhang, Z., Li, W., Xie, Y., Li, H. (2023). Risk stratification algorithm for clinical outcomes in anemic patients undergoing percutaneous coronary intervention. Annals of Medicine, 55(2): 2249200. https://doi.org/10.1080/07853890.2023.2249200

[24] Wang, X., Li, C., Xu, L., Jiang, S., Guan, C., Che, L., Wang, Y., Man, X., Xu, Y. (2025). Construction and validation of prognostic models for acute kidney disease and mortality in patients at risk of malnutrition: An interpretable machine learning approach. Clinical Kidney Journal, 18(4): sfaf080. https://doi.org/10.1093/ckj/sfaf080

[25] Mohammadi, I., Rajai Firouzabadi, S., Hosseinpour, M., Akhlaghpasand, M., Hajikarimloo, B., Zeraatian-Nejad, S., Sardari Nia, P. (2024). Using artificial intelligence to predict post-operative outcomes in congenital heart surgeries: A systematic review. BMC Cardiovascular Disorders, 24: 718. https://doi.org/10.1186/s12872-024-04336-6

[26] Elsamadicy, A.A., Koo, A.B., Reeves, B.C., Cross, J.L., Hersh, A., Hengartner, A.C., Karhade, A.V., Pennington, Z., Akinduro, O.O., Lo, S.F.L., Gokaslan, Z.L., Shin, J.H., Mendel, E., Sciubba, D.M. (2024). Utilization of machine learning to model important features of 30-day readmissions following surgery for metastatic spinal column tumors: The influence of frailty. Global Spine Journal, 14(4): 1227-1237. https://doi.org/10.1177/21925682221138053

[27] Mirallas, O., Martin-Cullell, B., Navarro, V., et al. (2024). Development of a prognostic model to predict 90-day mortality in hospitalised cancer patients (PROMISE tool): A prospective observational study. The Lancet Regional Health – Europe, 46: 101063. https://doi.org/10.1016/j.lanepe.2024.101063

[28] Mohammadi, S.M., Rumyantsev, M., Abdeeva, E., Baimukhambetova, D., et al. (2025). Post-COVID-19 condition prediction in hospitalised cancer patients: A machine learning-based approach. Cancers, 17(4): 687. https://doi.org/10.3390/cancers17040687