DOI: 10.53389/RJAHS.2026050112

OPEN ACCESS

Original Article

DOI: 10.53389/RJAHS.2026050112

Evaluating the appropriateness of ChatGPT in providing dietary advice: An evidence-based approach

Ayesha Saeed[1*], Maliha Amanat[1], Ayesha Ishafaq Ali[1]

[1]Department of Nutrition Sciences, University of South Asia, Lahore, Punjab, Pakistan.

Author`s Contribution

[AS] Conception and design, [MA, AI] Analysis and interpretation of the data, [MA, AI] Statistical expertise, [AS, MA, AI] Final approval and guarantor of the article

Article Info.

Received: March 28, 2026
Acceptance: May 13, 2026

Conflict of Interest: None
Funding Sources: None

Address of Correspondence

Ayesha Saeed
Department of Nutrition Sciences,
University of South Asia
Email: ayesha.saeed@usa.edu.pk

ABSTRACT

Background: ChatGPT, a generative artificial intelligence chatbot developed by OpenAI, is being increasingly used to obtain health and nutrition information. Limited research has evaluated the appropriateness and validity of dietary advice.

Objective: This study aimed to evaluate the appropriateness of the dietary advice provided by ChatGPT in alignment with international clinical guidelines for the management of seven specific non-communicable diseases, and its appropriateness in managing complex cases.

Methodology: Cross-sectional descriptive research was conducted in December 2025-January 2026. Seven NCD's were selected: obesity, dyslipidemia, non-alcoholic fatty liver disease (NAFLD), type 2 diabetes mellitus (T2DM), arterial hypertension, sarcopenia, and chronic kidney disease (CKD). For each disease, prompts were given to ChatGPT version 5.0 in independent sessions. Responses were then compared with international practice guidelines. Two experienced dietitians independently assessed each response using a structured evaluation framework. Any discrepancies between the dietitians were reviewed and finalised by a third reviewer, who made the final decision. In addition, a complex case scenario involving T2DM, obesity, and CKD was used to evaluate ChatGPT's ability to integrate recommendations from different clinical nutrition guidelines. Frequencies and percentages were calculated using SPSS v. 26.

Results: ChatGPT generally provided appropriate advice for individual diseases, dyslipidemia (88.8%), arterial hypertension (83.3%), obesity (93.8%), T2DM (94.1%), NAFLD (95.8%), and sarcopenia (94.4%). However, for CKD, 70% responses were appropriate, and six key recommendations were missing. ChatGPT demonstrated strong clinical reasoning when managing complex scenarios. AI maintained safe boundaries by explicitly stating it is not a medical professional and strongly advising users to seek highly individualized care from a registered renal dietitian.

Conclusion: ChatGPT provided largely appropriate and safe dietary advice, with good performance in both single-disease and complex clinical scenarios, though greater caution is needed in conditions requiring highly specific restrictions such as CKD.

Key words: Generative Artificial Intelligence, Communicable Diseases, Practice Guidelines, Nutritionists

Introduction

In recent years, artificial intelligence technologies have become increasingly prominent in healthcare communication. One such technology is ChatGPT, developed by OpenAI in November 2022, which functions as a large language model which is capable of generating detailed responses to user queries. By processing user prompts through machine-learning algorithms, the system produces structured and contextually relevant information while admitting its mistakes, and rejecting inappropriate requests.[1]

Recent research assessed the appropriateness of ChatGPT in generating nutritionally adequate, safe, and individualized diet plans for hypothetical individuals with non-communicable chronic diseases in comparison to a Registered Dietitian. The authors concluded that ChatGPT diet plans lacked personalization and often failed to address specific nutritional needs. In several cases, contraindicated foods were included.

They noted ethical and safety risks of using AI tools in clinical nutrition without supervision of a healthcare provider. [2]

A scoping review assessed the use of AI in supporting dietitians in the stages of the nutrition care process. Out of ninety-seven studies, seventy-nine studies reported patient-related outcomes, with mixed findings. Higher patient engagement was observed with Chatbots. Seventeen studies raised concerns encompassing ethics and patient safety. They concluded that AI education should be integrated into dietetic training and professional development to ensure safe and effective use in practice.[3]

With a dual burden of malnutrition-under-nutrition and over-nutrition-alongside a strained healthcare infrastructure, the demand for accessible dietary counseling far exceeds the available supply of trained dietitians and nutritionists. This disparity creates a critical gap that digital health tools like ChatGPT could potentially address. However, internet connectivity and digital literacy remains a barrier to the effective use of AI-based tools among certain segments of the population. This study aimed to evaluate the appropriateness of the dietary advice provided by ChatGPT in alignment with international clinical guidelines for the management of seven specific non-communicable diseases, and its appropriateness in managing complex cases.

Materials and Methods

Study design and Settings:

A cross-sectional descriptive research design was used. The study was conducted virtually, using ChatGPT (developed by OpenAI, version 5.0). The dietary advice was generated in December 2025 and evaluated by dietitians in January 2026.

Sample Size and Sampling Technique

The sampling technique was non-probability purposive sampling. The researchers deliberately selected a set of specific NCD's. The expert panel (two dietitians and an independent reviewer) was purposefully selected to evaluate the outputs based on their expertise.

Dietitian and Independent Reviewer

The assessors had 5-15 years of experience in healthcare or academia. Dietitian A had an M.Phil. & PGD in Dietetics, with 15 years of experience as a consultant dietitian. Dietitian B had an MS and PGD in Dietetics, with 12 years of experience as a consultant dietitian. The independent reviewer had an M.Phil. & PGD in Dietetics with 5 years of teaching experience.

Experiment 1:

Seven diseases were included: dyslipidemia (hypercholesterolemia and hypertriglyceridemia), arterial hypertension, type 2 diabetes mellitus (T2DM), obesity, non-alcoholic fatty liver disease (NAFLD), chronic kidney disease (CKD), and sarcopenia.

Experiment 2:

Complex Case Scenario (CKD, Obesity, T2DM).

Types of Variables

The set of prompts and Clinical guidelines were independent variables (qualitative variables).

The dependent variables were ChatGPT responses (qualitative variable) and expert evaluation of the responses using the evaluation rubric (qualitative and categorical variables).

Outcome Measure Tools

Socio-Demographic Questionnaire: The dietitians and third reviewer were asked about their academic qualifications, years of experience, clinical expertise, and current professional affiliation.

Clinical Guidelines: Clinical guidelines for dyslipidemia and hypertriglyceridemia, [4,5] arterial hypertension, [6,7] obesity, [8, 9] T2DM, [10,11] NAFLD, [12,13] CKD, [14,15] and sarcopenia [16,17] were employed as standards to evaluate and compare responses of ChatGPT.

Evaluation Rubric: Dietitians assessed the appropriateness of responses through an Evaluation rubric. Responses were classified as 6=Appropriate, 5=Not fully Matched, 4=General Advice, 3= Not Supported, 2=Inappropriate, 1=Missing. [18]

Reliability and validity: Although the Evaluation rubric was not a standardized instrument, its content validity was supported by using evidence-based clinical guidelines as the evaluation benchmark. Reliability was ensured through independent assessments by multiple experts and by calculating Cronbach's alpha and interclass correlations (Table 2).

Data Collection Procedure:

For the recruitment of dietitians, data collection letters were issued by the Department of Nutrition Sciences at the University of South Asia, Lahore.

Experiment 1: In the first step, ChatGPT version 5.0 was used in December 2025. Seven medical conditions were selected, namely, obesity, dyslipidemia, NAFLD, T2DM, arterial hypertension, sarcopenia, and CKD.

For each condition, three structured prompt types were used. Could you provide guidance on planning an optimal diet to

manage [disease]? What are the dietary recommendations for [disease]? and I have [disease] what should I eat? [18]

Quality Assurance: Each prompt was submitted three times in new sessions. New chat sessions were initiated for each prompt to avoid memory bias. The most comprehensive response (that answered all three questions and included maximum detail) for each prompt was selected for analysis. Responses were evaluated against condition-specific international dietary guidelines.

Assessment: Two experienced dietitians independently rated ChatGPT's answers using a standardized rubric. They independently evaluated the chatbot's answers using predefined categories. Appropriate: If the content was in accordance with the guidelines; Not Fully Matched: If it did not fully fulfill the recommendations of the guidelines; General Advice: Each advice that was not specifically focused but rather promoted overall a healthy diet; Not Supported: If did not find confirmation in the guidelines; Inappropriate: If contradicted the guidelines; Missing: Any dietary guideline recommendation that was not present in the response.[18] Disagreements were resolved by a third reviewer.

Experiment 2 (Complex Scenario): A hypothetical patient with T2DM, obesity, and CKD was used to assess ChatGPT's ability to integrate multiple guidelines. Responses were analyzed for contradictions or omissions. Prompt employed was: I have type 2 diabetes mellitus, obesity, and chronic kidney disease. Can you give me nutritional advice? [18]

Statistical Analysis:

SPSS version 26 was used to calculate the descriptive statistics of the responses of appropriateness of ChatGPT's dietary advice across different NCDs. The response was considered appropriate overall if the appropriate category exceeded 80% (Table 1). In Table 3 ChatGPT's dietary advice and foods to avoid were totaled for each disease. Percentages were calculated for all categories. For example, for hypercholesterolemia, the total ChatGPT responses were 18. The third reviewer classified 16 as appropriate and 2 as missing. The percentage of appropriate was calculated as 16/18x100=88.8%.

Ethical Considerations:

The research was conducted in accordance with the ethical standards of the Institutional Review Board, Faculty of Allied Health Sciences, University of South Asia (IRB No.: IRB-USA-FAHS/2025/1035). Anonymity and confidentiality of the dietitians and the third reviewer were ensured and acknowledged in the publication. Participation was voluntary, and informed consent was obtained.

Results

Table 1 Comparison of Assessment of Dietitians on responses from ChatGPT

Disease	Dietitian A	Dietitian B	Third Reviewer
Hypercholesterolemia	Appropriate	Appropriate	Appropriate
Hypertriglyceridemia	Appropriate	Appropriate	Appropriate
Hypertension	Appropriate/GA*	Appropriate	Appropriate
Obesity	Appropriate	Appropriate	Appropriate
T2DM**	Appropriate	Appropriate	Appropriate
NAFLD***	Appropriate/GA*	Appropriate	Appropriate
CKD****	Not Fully Matched	Not Fully Matched	Inappropriate
Sarcopenia	Appropriate	Appropriate	Appropriate

*GA General advice (one) **Type II Diabetes Mellitus ***Non-Alcoholic Fatty Liver Disease ****chronic kidney disease

The response was considered appropriate overall if the appropriate category exceeded 80%. Table 1 shows that dietary advice by ChatGPT was considered appropriate for hypercholesterolemia, hypertriglyceridemia, obesity, T2DM, and sarcopenia by all three evaluators. Hypertension and NAFLD were considered appropriate with one general advice, but overall scored appropriate by the third reviewer. CKD had the most inappropriate responses classified as not fully matched by the two independent dietitians.

Table 2 Cronbach Alpha and Interclass Correlation of Dietitians and Third Reviewer

Intraclass Correlation		95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	0.833	0.554	0.960	16.000	7	14	0.000
Average Measures	0.938	0.789	0.986	16.000	7	14	0.000

The intraclass correlation coefficient (ICC) showed a high level of agreement between the dietitians and the third reviewer. The

single-measures ICC was 0.833 (95% CI: 0.554–0.960, p < 0.001), indicating good reliability of the ratings provided by an individual reviewer. The average ICC was 0.938 (95% CI: 0.789–0.986; p < 0.001), indicating excellent reliability when reviewer ratings were averaged. The statistically significant F-test further confirms that the observed agreement was not due to chance. Cronbach's alpha was 0.938, indicating excellent internal consistency.

Table 3 Classification of ChatGPT Responses According to Evaluation Rubric

	6 N (%)	5 N (%)	4 N (%)	3 N (%)	2 N (%)	1 N (%)
Hyper-cholesterolemia	16(88.8 %)	0(0 %)	0(0%)	0(0%)	0(0%)	2(11.1 %)
Hyper-triglyceridemia	16(88.8 %)	0(0 %)	0(0%)	0(0%)	0(0%)	2(11.1 %)
Arterial Hypertension	10(83.3 %)	0(0 %)	0(0%)	0(0%)	1(8.3 %)	1(8.3 %)
Obesity	15(93.8 %)	0(0 %)	1(6.2 %)	0(0%)	0(0%)	0(0%)
T2DM	16(94.1 %)	0(0 %)	0(0%)	0(0%)	0(0%)	1(5.9 %)
CKD	14 (70%)	0(0 %)	0(0%)	0(0%)	0(0%)	6 (30%)
NAFLD	23 (95.8%)	0(0 %)	0(0%)	0(0%)	0(0%)	1(4.2 %)
Sarcopenia	17(94.4 %)	0(0 %)	0(0%)	1(5.6 %)	0(0%)	0(0%)

6=Appropriate, 5=Not fully Matched, 4=General Advice, 3= Not Supported, 2=Inappropriate, 1=Missing

Table 3 highlights that mixed results were obtained across the seven diseases. The appropriateness for CKD was the lowest with an appropriate ratio of 70% up to 95.8% for NAFLD. If we consider all the correct recommendations, including both "appropriate" and "general advice", the overall appropriateness of nutritional recommendations reached high concordance for Obesity.

Table 4 identifies the significant gaps in clinical precision as per assessment of the third reviewer. The most notable deficiencies occurred in CKD, where six key recommendations were missing, including exact ranges for potassium and phosphorus. Furthermore, the model provided inappropriate advice for NAFLD (recommending "lots of fresh fruits") and Hypertension (failing to specify salt and potassium targets).

Table 4 Inappropriate and Missing Recommendations of Third Reviewer

Types	Responses	
Hypercholesterolemia	Missing (2)	Consume 25-30 g fiber/ day Avoid sugary food items
Hypertriglyceridemia	Missing (2)	Consume 25-30 g fiber/ day Avoid sugary food items
Arterial Hypertension	Inappropriate (1) Missing (1)	Reduce salt intake to <5g/ day Exact range of potassium-rich foods
Obesity	General advice (1)	Use olive oil / canola oil, avoid ghee, butter, and fried foods
T2DM	Missing (1)	Improve QOL, manage stress and sleep
CKD	Missing (6)	Exact range of potassium-rich foods Exact range for phosphorus intake Avoid green leafy vegetables and spinach Avoid Bakery food items Avoid Carbonated beverages Intake range for salt substitute
NAFLD	Inappropriate (1)	Eat lots of fresh fruits
Sarcopenia	Not Supported (1)	Exact range of protein intake

Complex Case Scenario A Conversation with ChatGPT

Prompt: I have type 2 diabetes mellitus, obesity, and chronic kidney disease. Can you give me nutritional advice?

Chat GPT Answer:

I am an AI, not a doctor or a registered dietitian. The information below is for educational purposes and should not replace professional medical advice. Because your health conditions require highly individualized care, it is crucial to consult your doctor or a registered renal dietitian before making any changes to your diet.

Balancing Type 2 diabetes, obesity, and chronic kidney disease (CKD) can be challenging because the standard advice for one condition sometimes conflicts with another. For example, a typical diabetes diet often encourages whole grains, beans, and certain fruits, but a CKD diet might restrict those exact same foods if they are high in potassium or phosphorus.

Because of this overlap, your nutritional plan needs to be carefully tailored to the specific stage of your kidney disease and your current lab results. Here is a general overview of how

nutrition is typically managed when these three conditions are present.

1. Managing Carbohydrates (Diabetes & Obesity)

To manage blood sugar and weight, the focus is usually on the quality and portion size of carbohydrates, while keeping kidney limits in mind.

Carb Counting: Spacing your carbohydrates evenly throughout the day helps prevent blood sugar spikes.

Kidney-Friendly Carbs: Some complex carbs (like brown rice or potatoes) are high in phosphorus or potassium. You may need to swap them for lower-potassium options like white rice, refined pasta, sourdough bread, or specific low-potassium fruits and vegetables (like berries, apples, or green beans).

2. Monitoring Protein (CKD)

Protein is essential, but damaged kidneys have a hard time filtering the waste products that come from breaking it down.

Quality over Quantity: Depending on your CKD stage, your doctor may recommend a moderate or reduced protein intake.

Plant-Based Proteins: Many nephrologists now recommend shifting toward plant-based proteins (like tofu or certain lentils, if potassium/phosphorus levels allow) because they generate fewer waste products and are easier on the kidneys than red meat.

3. Watching Minerals (CKD)

As CKD progresses, your kidneys may struggle to filter out excess minerals from your blood. Your doctor will monitor your labs to tell you if you need to restrict these:

Sodium (Salt): High sodium increases blood pressure, which causes further damage to the kidneys. Avoid processed foods, canned soups, and fast food. Cook with herbs and spices instead of salt.

Potassium: If your potassium levels are high, you will need to limit foods like bananas, tomatoes, potatoes, oranges, and avocados.

Phosphorus: High phosphorus can pull calcium from your bones. You may need to limit dairy products, dark colas, nuts, and processed meats (which often contain added phosphorus preservatives).

4. Calorie & Fat Management (Obesity & Diabetes)

Gradual, sustainable weight loss can significantly improve insulin sensitivity and reduce the workload on your kidneys.

Heart-Healthy Fats: People with diabetes and CKD are at a higher risk for heart disease. Focus on healthy fats like olive oil and avoid trans fats and saturated fats.

Portion Control: Even healthy fats are calorie-dense. Managing your overall portion sizes is the safest way to encourage gradual weight loss without depriving your body of nutrients.

5. Hydration

Fluid intake is highly individualized. While staying hydrated is generally good, advanced CKD may require fluid restriction to prevent swelling and fluid buildup in the lungs. Always follow your doctor's specific guidance on how much water you should drink daily.

The Most Important Step: Because managing these three conditions simultaneously is complex, the best thing you can do is ask your nephrologist or primary care doctor for a referral to a registered dietitian who specializes in renal (kidney) nutrition. They can look at your recent blood work (A1C, GFR, potassium, phosphorus) and create a specific meal plan just for you.

Discussion

The present study evaluated the appropriateness of ChatGPT in providing dietary advice for seven non-communicable diseases and one complex case scenario by comparing its responses with international evidence-based dietary guidelines and expert review. In experiment 1, overall, the findings indicate that ChatGPT performed well for most conditions, with advice being judged appropriate for hypercholesterolemia, hypertriglyceridemia, obesity, type 2 diabetes mellitus (T2DM), and sarcopenia by all three evaluators. Responses for hypertension and non-alcoholic fatty liver disease (NAFLD) were also largely acceptable, although some advice was categorized as general rather than fully guideline-specific. In contrast, chronic kidney disease (CKD) showed the poorest performance, with responses rated as not fully matched by two dietitians and inappropriate by the third reviewer.

These findings are broadly consistent with Ponzo et al., [18] who reported that ChatGPT can provide generally acceptable dietary advice, but that its performance varies depending on the clinical scenario and the level of specificity required. Similar to the current study, Ponzo et al. [18] observed that ChatGPT was more reliable in delivering broad healthy eating recommendations than in providing nuanced, disease-specific nutritional prescriptions. The strong performance seen here in hypercholesterolemia and hypertriglyceridemia may be because recommendations for lipid disorders are relatively well established and often center on general dietary principles such

as increasing fiber intake, reducing saturated fats, and limiting sugars, as emphasized in ESC/EAS and AHA/ACC cholesterol guidelines. [4,5] Because these recommendations are common across many educational and clinical sources, ChatGPT may have been more likely to reproduce them appropriately.

The high appropriateness observed for obesity and T2DM may likewise reflect the broad availability and consistency of medical nutrition therapy recommendations for these conditions. Current obesity guidelines emphasize calorie control, healthier fat choices, reduction of ultra-processed and fried foods, and sustainable dietary behavior change, [8,9] while diabetes guidelines highlight individualized meal planning, improved dietary quality, lifestyle behavior, and long-term self-management. [10,11] In the present study, obesity achieved particularly high appropriateness, and when "appropriate" and "general advice" were considered together, the recommendations reached full overall acceptability. This suggests that ChatGPT can successfully generate practical lifestyle-oriented advice for disorders in which nutrition counseling often begins with general but evidence-based principles. However, one weakness remained: even in obesity, some advice was generic rather than sufficiently detailed, showing that acceptable guidance is not always equivalent to complete guideline adherence.

For hypertension, the findings were somewhat more mixed. Although the overall judgment remained appropriate, the model failed to specify important quantitative targets such as reducing salt intake to less than 5 g/day and providing exact potassium-related guidance. This is relevant because hypertension dietary management depends not only on general advice to "eat less salt" or consume more potassium-rich foods, but also on explicit and clinically actionable targets, as highlighted by both the European Society of Hypertension and International Society of Hypertension guidelines. [6,7] Therefore, ChatGPT's response may have appeared reasonable on the surface while still lacking the precision needed for optimal dietary counseling. This partially agrees with Ponzo et al., [18] who also suggested that large language models may provide plausible answers that miss critical technical details.

The result for NAFLD was similarly strong overall, with a very high proportion of appropriate responses, but one recommendation was considered inappropriate because it advised eating "lots of fresh fruits." While fruit intake is generally part of a healthy dietary pattern, NAFLD nutrition management requires balanced, context-specific advice focused on weight reduction, avoidance of excess fructose and energy intake, and overall dietary quality. [12,13] In this sense, discrepancy may reflect the tendency of AI-generated responses to frame advice in overly simplified positive terms without considering metabolic nuances. This does not mean the entire response was poor, but it highlights a limitation in translating general healthy eating principles into disease-specific metabolic guidance.

The most important finding of this study was ChatGPT's weak performance for CKD. Compared with the other disease categories, CKD had the lowest proportion of appropriate recommendations and the highest number of missing elements. Missing advice included exact potassium and phosphorus ranges, restrictions on certain foods, avoidance of carbonated beverages and bakery products, and intake guidance related to salt substitutes. This result is understandable because CKD nutrition therapy is one of the most complex areas of clinical dietetics, requiring individualized recommendations based on disease stage, biochemical profile, and risk of electrolyte imbalance. [14,15] Unlike obesity or dyslipidemia, CKD management depends heavily on precise quantitative adjustments and careful clinical interpretation. As a result, a general-purpose language model may struggle to produce complete and safe recommendations without condition-specific prompting or contextual clinical data. This finding is strongly aligned with Ponzo et al., [18] who noted that while ChatGPT can be useful for general advice, it should not be relied upon for complex or high-risk medical nutrition decisions.

The sarcopenia findings were largely favorable, with most responses rated appropriate, although one response was considered not supported because it did not provide an exact protein intake range. Current sarcopenia guidance emphasizes adequate protein intake alongside exercise, especially resistance training, for preservation of muscle mass and function. [16,17] The generally good performance may reflect that ChatGPT correctly captured the broad conceptual framework of sarcopenia management. However, the absence of a precise protein target again illustrates a recurring limitation: ChatGPT often performs better at identifying the direction of advice than at supplying exact therapeutic thresholds.

The differences in appropriateness likely stem from the generalized nature of Large Language Models (LLMs). While standard guidelines for obesity or T2DM translate easily into broad, actionable advice, conditions like CKD require highly individualized, precise quantitative targets (e.g., specific milligram limits of phosphorus) based on continuous lab monitoring. AI models prioritize safe, generalized text over definitive, prescriptive clinical limits to avoid medical liability.

For hypercholesterolemia, hypertriglyceridemia, obesity, T2DM, NAFLD, and sarcopenia, ChatGPT's advice was largely consistent with guideline-based recommendations. However,

for CKD, substantial omissions and inappropriate advice were identified. Therefore, the findings indicate that ChatGPT does not consistently match international dietary guidelines across all seven diseases.

In experiment 2, when ChatGPT was presented with the complex overlapping conditions of T2DM, obesity, and CKD, ChatGPT successfully identified the conflicting nature of standard dietary advice such as whole grains being beneficial for diabetes but potentially harmful for CKD due to high potassium or phosphorus. The AI correctly prioritized safety by explicitly stating it is not a doctor, providing information for educational purposes only, and strongly recommending consultation with a registered renal dietitian.

ChatGPT's reasoning in this complex scenario complied with international guidelines. Its emphasis on tailoring carbohydrate choices for both glycemic control and renal safety align perfectly with the intersections of ADA diabetes standards [10] and CKD nutrition guidelines. [14] Furthermore, its recommendation to shift toward plant-based proteins, provided potassium and phosphorus levels allow, reflects modern nephrology practices outlined in updated clinical guidelines. [15] The AI's approach to caloric management and heart-healthy fats for weight loss is consistent with Canadian Adult Obesity Clinical Practice Guidelines. [9] By defaulting to individualized care and referring out to specialists, the model adhered to the ethical and practical standards expected of clinical practitioners.

An important strength of this study is that the evaluation was based on international clinical guidelines and involved independent review by two dietitians with adjudication by a third reviewer. This strengthens the objectivity and clinical relevance of the assessment. In addition, the excellent average-measures intraclass correlation and strong single-measure ICC indicate a high level of agreement among reviewers, supporting the reliability of the evaluation process. Another strength is that the study assessed multiple disease conditions rather than a single diagnosis, allowing a broader understanding of where ChatGPT performs well and where it is less reliable.

Earlier studies [2,3,18] had highlighted limitations, ethical and safety concerns, of earlier versions of ChatGPT (version 3.5 and 4.0), and identified the need for continuous evaluation for dietary advice, dietary planning, and application in the nutrition care process. This study contributes important evidence on the clinical appropriateness and validity of ChatGPT version 5.0 for dietary advice across major non-communicable diseases and complex case scenarios. The study supports the idea that ChatGPT may serve as a supplementary educational tool, but not as a replacement for professional dietetic judgment in complex disease management.

Limitations:

First, it assessed a limited number of disease scenarios and may not capture the full variability of ChatGPT responses across different prompts or patient contexts. Second, the evaluation focused on appropriateness against guidelines but did not measure patient understanding, usability, or clinical outcomes. Third, ChatGPT responses may vary over time and across model versions, which can affect reproducibility. Finally, because CKD and some other conditions require individualized clinical interpretation, a single static prompt may not fully test the model's capacity when more detailed patient information is provided.

Conclusion

The evaluation indicates that ChatGPT is highly capable of providing appropriate dietary advice for several non-communicable diseases, achieving up to high appropriateness for conditions like obesity. Reviewers demonstrated excellent consistency and agreement in these positive assessments. However, the AI struggles significantly with conditions that demand strict, quantitative clinical parameters like chronic kidney disease. ChatGPT demonstrated strong clinical reasoning when managing complex scenarios. This study contributes important evidence on the clinical appropriateness and validity of ChatGPT version 5.0.

Recommendations:

1. Use ChatGPT for nutrition advice, not in place of a dietitian.

2. Always confirm the AI advice, in case of complex conditions and multiple illnesses.

3. Ask detailed and specific questions in the prompt, for more appropriate responses.

4. Use ChatGPT to improve health awareness and prepare informed questions for healthcare professionals.

Acknowledgement:

We would like to extend special acknowledgement to our dietitians (Dietitian A: RD Ms. Shifa Amna Ali, Dietitian B: RD Ms. Hajra Saleem) and third reviewer (Assistant Professor, Ms. Taliya Amir) for their invaluable guidance, expert insight, and significant contribution to this project.

References

1. OpenAI. Introducing ChatGPT. OpenAI. November 30, 2022. Accessed May 29, 2026. URL https://openai.com/index/chatgpt/

2. Onay T, Bekar D, Çoban E, Doğan N, Günşen U. Artificial Intelligence in Clinical Nutrition: A Descriptive Comparison of ChatGPT-and Dietitian-Planned Diets for Chronic Disease Scenarios. J Hum Nutr Diet. 2025; 38(5):e70135. https://doi.org/10.1111/jhn.70135

3. Ngo K, Mekhail S, Chan V, Li X, Yin A, Choi HY, Allman-Farinelli M, Chen J. The use of artificial intelligence (AI) to support dietetic practice across primary care: a scoping review of the literature. Nutrients. 2025; 17(22):3515. https://doi.org/10.3390/nu17223515
4. Mach F, Baigent C, Catapano AL, Koskinas KC, Casula M, Badimon L, Chapman MJ, De Backer GG, Delgado V, Ference BA, Graham IM. 2019 ESC/EAS Guidelines for the management of dyslipidaemias: Lipid modification to reduce cardiovascular risk: The task force for the management of dyslipidaemias of the European Society of Cardiology (ESC) and European Atherosclerosis Society (EAS). Eur Heart J. 2020; 41(1):111-88. https://doi.org/10.1093/eurheartj/ehz455
5. Grundy SM, Stone NJ, Bailey AL, Beam C, Birtcher KK, Blumenthal RS, Braun LT, De Ferranti S, Faiella-Tommasino J, Forman DE, Goldberg R. 2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA Guideline on the management of blood cholesterol: A report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. J Am Coll Cardiol. 2019; 73(24):e285-350. https://doi.org/10.1016/j.jacc.2018.11.003
6. Mancia G, Kreutz R, Brunström M, Burnier M, Grassi G, Januszewicz A, Muiesan ML, Tsioufis K, Agabiti-Rosei E, Algharably EA, Azizi M. 2023 ESH Guidelines for the management of arterial hypertension. The Task Force for the management of arterial hypertension of the European Society of Hypertension: Endorsed by the International Society of Hypertension (ISH) and the European Renal Association (ERA). J Hypertens. 2023; 41(12):1874-2071. https://doi.org/10.1097/HJH.0000000000003480
7. Unger T, Borghi C, Charchar F, Khan NA, Poulter NR, Prabhakaran D, Ramirez A, Schlaich M, Stergiou GS, Tomaszewski M, Wainford RD. 2020 International Society of Hypertension global hypertension practice guidelines. Hypertension. 2020; 75(6):1334-57. https://doi.org/10.1161/hypertensionaha.120.15026
8. Yumuk V, Tsigos C, Fried M, Schindler K, Busetto L, Micic D, Toplak H. European guidelines for obesity management in adults. Obes Facts. 2015; 8(6):402-24. https://doi.org/10.1159/000442721
9. Brown J, Clarke C, Johnson Stoklossa C, Sievenpiper J. Canadian Adult Obesity Clinical Practice Guidelines: Medical Nutrition Therapy in Obesity Management. Last updated: October 21, 2022. Available from: https://obesitycanada.ca/guidelines/nutrition. Accessed January 5, 2026.
10. ElSayed NA, Aleppo G, Aroda VR, Bannuru RR, Brown FM, Bruemmer D, Collins BS, Hilliard ME, Isaacs D, Johnson EL, Kahan S. 5. Facilitating positive health behaviors and well-being to improve health outcomes: Standards of care in diabetes-2023. Diabetes care. 2023; 46(Supplement 1):S68-96. https://doi.org/10.2337/dc23-S005
11. The Diabetes and Nutrition Study Group (DNSG) of the European Association for the Study of Diabetes (EASD). Evidence-based European recommendations for the dietary management of diabetes. Diabetologia. 2023; 66:965-985. https://doi.org/10.1007/s00125-023-05894-8
12. Rinella ME, Neuschwander-Tetri BA, Siddiqui MS, Abdelmalek MF, Caldwell S, Barb D, Kleiner DE, Loomba R. AASLD Practice Guidance on the clinical assessment and management of nonalcoholic fatty liver disease. Hepatology. 2023; 77(5):1797-835. https://doi.org/10.1097/HEP.0000000000000323
13. Bischoff SC, Bernal W, Dasarathy S, Merli M, Plank LD, Schütz T, Plauth M. ESPEN practical guideline: Clinical nutrition in liver disease. Clin Nutr. 2020; 39(12):3533-62. https://doi.org/10.1016/j.clnu.2020.09.001
14. Ikizler TA, Burrowes JD, Byham-Gray LD, Campbell KL, Carrero JJ, Chan W, Fouque D, Friedman AN, Ghaddar S, Goldstein-Fuchs DJ, Kaysen GA. KDOQI clinical practice guideline for nutrition in CKD: 2020 update. Am J Kidney Dis. 2020; 76(3): S1-07. https://doi.org/10.1053/j.ajkd.2020.05.006.
15. Levin A, Ahmed SB, Carrero JJ, Foster B, Francis A, Hall RK, Herrington WG, Hill G, Inker LA, Kazancıoğlu R, Lamb E. Executive summary of the KDIGO 2024 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease: known knowns and known unknowns. Kidney Int. 2024;105(4):684-701. Available at: https://kdigo.org/wp-content/uploads/2017/02/KDIGO-2024-CKD-Guideline-Executive-Summary.pdf
16. Deutz NE, Bauer JM, Barazzoni R, Biolo G, Boirie Y, Bosy-Westphal A, Cederholm T, Cruz-Jentoft A, Krznariç Z, Nair KS, Singer P. Protein intake and exercise for optimal muscle function with aging: recommendations from the ESPEN Expert Group. Clin Nutr. 2014; 33(6):929-36. https://doi.org/10.1016/j.clnu.2014.04.007
17. Dent E, Morley JE, Cruz-Jentoft AJ, Arai H, Kritchevsky SB, Guralnik J, Bauer JM, Pahor M, Clark BC, Cesari M, Ruiz J. International clinical practice guidelines for sarcopenia (ICFSR): screening, diagnosis and management. J Nutr Health Aging. 2018; 22(10):1148-61. https://doi.org/10.1007/s12603-018-1139-9
18. Ponzo V, Goitre I, Favaro E, Merlo FD, Mancino MV, Riso S, Bo S. Is ChatGPT an effective tool for providing dietary advice?. Nutrients. 2024;16(4):469. https://doi.org/10.3390/nu16040469