European Journal of Clinical Nutrition www.nature.com/ejcn

ARTICLE

Check for updates

Artificial intelligence chatbots for the nutrition management of diabetes and the metabolic syndrome

Farah Naja[1,2], Mandy Taktouk[2], Dana Matbouli[2], Sharfa Khaleel[1], Ayah Maher[1], Berna Uzun[3], Maryam Alameddine[4] and Lara Nasreddine[2]✉

BACKGROUND: Recently, there has been a growing interest in exploring AI-driven chatbots, such as ChatGPT, as a resource for disease management and education.

OBJECTIVE: The study aims to evaluate ChatGPT's accuracy and quality/clarity in providing nutritional management for Type 2 Diabetes (T2DM), the Metabolic syndrome (MetS) and its components, in accordance with the Academy of Nutrition and Dietetics' guidelines.

METHODS: Three nutrition management-related domains were considered: (1) Dietary management, (2) Nutrition care process (NCP) and (3) Menu planning (1500 kcal). A total of 63 prompts were used. Two experienced dietitians evaluated the chatbot output's concordance with the guidelines.

RESULTS: Both dietitians provided similar assessments for most conditions examined in the study. Gaps in the ChatGPT-derived outputs were identified and included weight loss recommendations, energy deficit, anthropometric assessment, specific nutrients of concern and the adoption of specific dietary interventions. Gaps in physical activity recommendations were also observed, highlighting ChatGPT's limitations in providing holistic lifestyle interventions. Within the NCP, the generated output provided incomplete examples of diagnostic documentation statements and had significant gaps in the monitoring and evaluation step. In the 1500 kcal one-day menus, the amounts of carbohydrates, fat, vitamin D and calcium were discordant with dietary recommendations. Regarding clarity, dietitians rated the output as either good or excellent.

CONCLUSION: Although ChatGPT is an increasingly available resource for practitioners, users are encouraged to consider the gaps identified in this study in the dietary management of T2DM and the MetS.

European Journal of Clinical Nutrition (2024) 78:887–896; https://doi.org/10.1038/s41430-024-01476-y

INTRODUCTION

Non-communicable diseases (NCDs) pose a significant public health challenge, being responsible for 74% of all deaths worldwide, in 2019 [1]. Among these NCDs, the World Health Organization (WHO) reported that diabetes caused 1.5 million deaths in 2019 [2], and, the Global Burden of Disease (GBD) study estimated that there were 529 million people living with diabetes in 2021 [3]. The GBD study also showed that the global age-standardized total diabetes prevalence was of 6.1%, with the highest prevalence rates being observed in Oceania (12.3%) and North Africa and the Middle East (9.3%) [3]. Importantly, the total diabetes prevalence essentially reflects type 2 diabetes (T2DM), which was found to account for 96% of total diabetes cases [3].

Metabolic risk factors of T2DM include the metabolic syndrome (MetS), characterized by central obesity, atherogenic dyslipidemia (high triglyceride levels or low levels of high-density lipoprotein, HDL), hyperglycemia, and hypertension (HTN) [4, 5]. MetS is usually defined as the co-presence of at least three of the five aforementioned metabolic abnormalities [4, 5]. A global meta-analysis showed a relatively high prevalence of MetS ranging between 12.5% to 31.4% [6]. Given its widespread prevalence and deleterious health implications, concerted efforts are needed to develop aggressive and contextualized public health interventions in addressing MetS and its risk factors. Available evidence showed that dietary interventions including adequate intakes of fruits, vegetables, whole grains and reduced consumption of simple sugars, saturated fats, trans fats, and cholesterol are paramount to prevent and manage T2DM and MetS [7]. The American Heart Association (AHA) and the American Diabetes Association (ADA) consider lifestyle modifications, including nutritional management, as the first-line therapy [7, 8], being 1.59 times more effective in reversing MetS compared to pharmaceutical treatments [9]. Healthcare professionals, particularly dietitians, play a key role in promoting dietary guidelines for T2DM, MetS and their complications. Easy access to reliable scientific information can facilitate the effective implementation of these guidelines in clinical practice.

Artificial Intelligence (AI) solutions that include AI-powered chatbots like the Chat Generative Pre-trained Transformer (ChatGPT), are being increasingly recognized for their applicability

[1]Department of Clinical Nutrition and Dietetics, College of Health Sciences, Research Institute of Medical and Health Sciences (RIMHS), University of Sharjah, Sharjah, United Arab Emirates. [2]Department of Nutrition and Food Sciences, Faculty of Agricultural and Food Sciences, American University of Beirut (AUB), Beirut, Lebanon. [3]Department of Mathematics, Near East University, Nicosia, Turkey. [4]Medical Program Royal College of Surgeons, Ireland, Manama, Bahrain. ✉email: ln10@aub.edu.lb

Received: 8 March 2024 Revised: 16 July 2024 Accepted: 17 July 2024
Published online: 26 July 2024

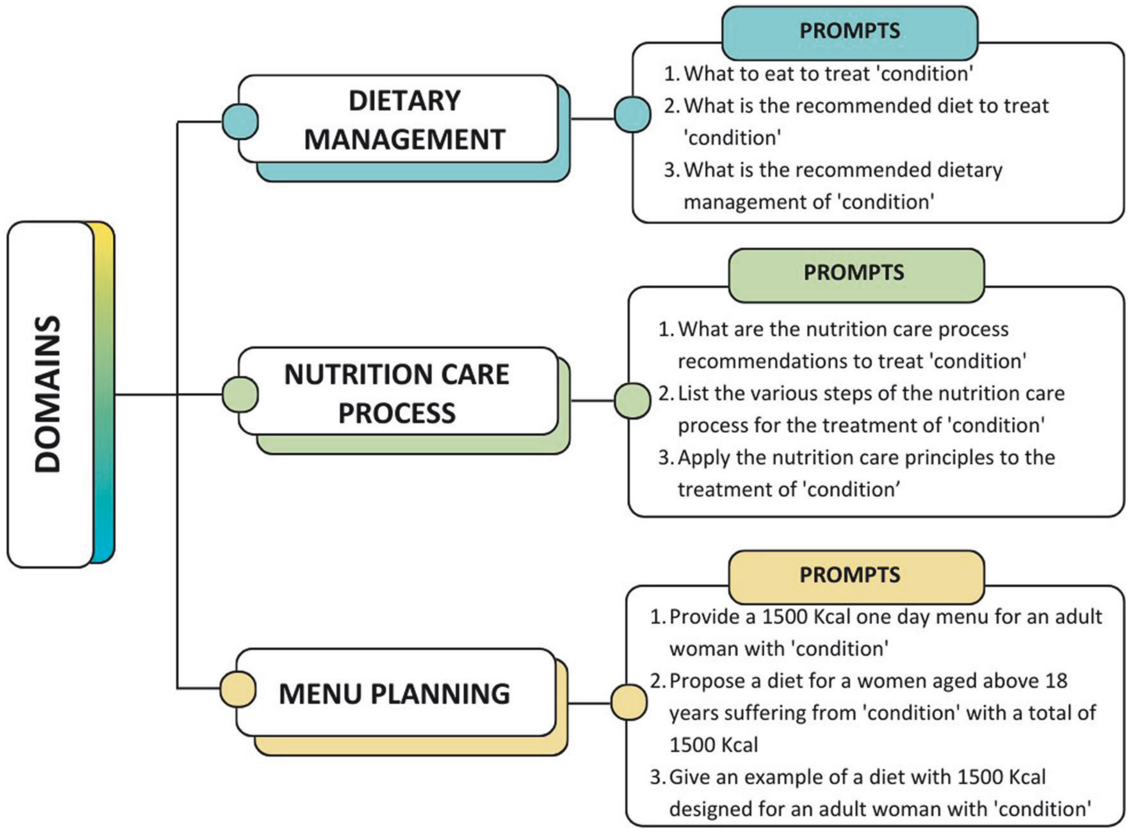

Fig. 1 The domains and corresponding prompts for T2DM, the MetS and its components, namely obesity, hyperglycemia (or prediabetes), HTN, low HDL levels, and hypertriglyceridemia.

in the management of chronic diseases, the latter requiring processing and synthesis of extensive diagnostics and treatment data [10]. ChatGPT, developed by OpenAI based on the GPT3 large language model, has gained unprecedented traction among healthcare professionals and the general public since its launch on November 30, 2022 [11]. Using deep learning, ChatGPT is pretrained on a large dataset of text stemming from online sources such as articles, books, websites, and other publicly available sources to provide human-like answers to text-based prompts [12]. Existing literature on the use of ChatGPT in healthcare predominantly addresses sciences like microbiology [13], nursing [14], parasitology [15], radiology [16], and medicine [17]. Although ChatGPT was considered as potentially promising in providing personalized recommendations related to obesity treatment, its use was also linked to several challenges including the lack of accountability in cases where AI-based models provide incorrect, incomplete or harmful advice [12, 18]. In addition, Chen et al. (2023) reported that AI-derived responses frequently mix incorrect recommendations among correct ones, an error difficult even for experts to detect [19]. The AI chatbot can also fabricate non-existent references [20] and generate "hallucinations" whereby output is nonsensical or unfaithful to the provided input or "prompt". The confident natural language used by conversational AI may lead users to overestimate the scientific rigor of the information presented [21].

In nutrition, ChatGPT's dynamic conversational capabilities offers potential for personalized and engaging education [18, 22–24]. An increasing number of real-world applications of ChatGPT have been launched in the field of nutrition and health, including nutritional counseling platforms, health and fitness apps, public health campaigns and chatbots' incorporations within school health and nutrition curricula [22, 25–28]. While these applications highlight the transformative potential of using ChatGPT in nutrition education and information, few research studies have explored its potential applications and limitations in providing clinical nutritional advice [29]. This study aims to evaluate ChatGPT's accuracy and quality/clarity to provide nutritional management for T2DM, the MetS and its components, in accordance with the guidelines of the Academy of Nutrition and Dietetics.

MATERIALS AND METHODS

Given the public health impact of T2DM and its risk factors, the diet-related metabolic diseases addressed in this study were T2DM, the MetS and its components, namely obesity (as a more general term for central obesity), hyperglycemia (or prediabetes), HTN, low levels of HDL, and hypertriglyceridemia (therefore 7 conditions in total) [4]. Three main nutrition management-related domains were considered for each of those conditions: (1) dietary management, (2) nutrition care process (NCP) recommendations and (3) menu planning. For each of those domains, three zero-shot prompt templates were designed to solicit recommendations. In designing the prompts for this study, the steps for prompt engineering to generate accurate, relevant, and useful responses, as recommended by Ray et al., 2023, were followed [30]: (1) Start with specific prompts. The dietetic or nutritional management of the condition was specified in all the prompts, using words such as 'eat', 'dietary', 'nutrition', and 'menu'; (2) Provide clear context. The disease or the condition was indicated in all prompts to provide the context of the nutrition and dietary guidelines. (3) Use open-ended questions. All prompts were kept open-ended including verbs such as 'list', 'provide', 'give an example' and (4) apply constraints. For example, in the 'menu planning' domain, constraints such as age 'adult', sex 'women' and amount of calories '1500 Kcal' were indicated. Following these steps, for every condition, 9 prompts were developed (three for each domain) (Fig. 1). The total number of prompts used for the 7 conditions was 63 prompts. The list of prompts used in this

Table 1. Evaluation of AI-derived dietary management of diabetes and the MetS.

Disease	Dietitian1[a]	Dietitian2[a]	Gaps
T2DM	1	2	Energy intake/expenditure, weight loss management, macronutrients intake, specific diet interventions (CHO-modified diet, protein-modified diet, fiber-modified diet, DASH, Mediterranean), recommended amounts/servings from each food group, CHO counting, glycemic load, timing of meals and snacks, hypoglycemia.
MetS	3	3	Energy intake, weight reduction.
Hyperglyemia (or prediabetes)	2	3	Meal planning and recommended amounts/serving from each food group.
Obesity	2	2	Energy deficit, macronutrients intake, nutrient-dense foods, lean and plant proteins, omega 3 fatty acids, SFA, added sugars, refined grains, intermittent fasting.
HTN	3	2	Healthy patterns with a focus on sodium and the different types of fat.
High TG	2	2	Energy intake, weight management, different types of fat (sources and intakes), soluble fiber, plant stanols/sterols, fish and fish oil, adequate and balanced food quantities.
Low HDL levels	1	2	Weight loss, healthy dietary patterns along with the servings and amounts from each food group, DASH/Mediterranean diet, different types of fat (sources, intakes and impact on health), viscous fiber, plants, DHA and EPA, Alphalinolenic acid, fish oil supplements.

AI artificial intelligence, *CHO* carbohydrates, *DASH* Dietary Approaches to Stop HTN, *DHA* docosahexaenoic acid, *EPA* eicosapentaenoic acid, *HDL* high-density lipoprotein, *HTN* hypertension, *MetS* metabolic syndrome, *SFA* saturated fatty acids, *T2DM* type 2 diabetes mellitus, *TG* triglycerides.
[a]1 = none; 2 = some but not all; 3 = most but not all; 4 = all.

study are found in Appendix A. These prompts were then fed into the GPT-3.5-turbo0301 model through the ChatGPT interface provided by OpenAI, during October 2023. Data collection ended 30 October 2023. Since the protocol of this study did not entail the active engagement or participation of any living individuals (or animals), the regulatory requirement for obtaining ethical approval, which typically governs studies involving human and animal subjects, was deemed unnecessary.

For the 'dietary management' and 'NCP' domains, the chatbot's recommendations were benchmarked against those indicated in the Nutrition Care Manual (NCM) ® of the Academy of Nutrition and Dietetics. These guidelines were selected for this study because the NCM is the main evidence-based, and up to date reference tool developed and managed by the Academy of Nutrition and Dietetics, the professional body for registered dietitians and nutrition specialists [31]. Additionally, the NCP, which is the core of the NCM, is a scientific problem-solving approach that ensures a structured method for identifying and addressing nutritional issues, thus standardizing assessment and intervention approaches, improving clinical outcomes and enhancing the effectiveness of nutritional management. The Academy works with a number of accreditation organizations, which further strengthens its role and outreach in various fields [30].

Two experienced dietitians evaluated the chatbot output's concordance with the guidelines. Both dietitians were licensed clinical practitioners. The main focus of the evaluation was on accuracy. The dietitians indicated their response using a Likert scale of 1 to 4, corresponding to 'not at all' (i.e., no concordance between the output and the guidelines), 'some but not all', 'most but not all' and 'all', respectively. 'Not all' was used when the ChatGPT output significantly diverged from the guidelines. 'Some but not all' was appropriate for outputs that partially aligned but had notable discrepancies. 'Most but not all' indicated substantial alignment with some deviations, while 'Al'l denoted a perfect match between the ChatGPT output and guidelines. The dietitians were trained to cross check the information obtained from the ChatGPT with the guidelines, and examine if there are any inconsistency or contradictions within the text. More specifically, three main areas were used in evaluating the accuracy of the output: relevance of the output, correctness of the information, and adherence of the output to the guidelines. For each of these areas, clear description corresponding to the various levels of the Likert scale used in the study was provided to the dietitians. Previous studies recommended the use of similar criteria in evaluating the medical responses of the ChatGPT [32, 33]. Appendix B presents specific criteria that were used by the dietitians to evaluate the outputs. For ratings 1, 2 or 3, the dietitians indicated the main gaps that they have observed. Training sessions were held to standardize the evaluation process and to ensure inter-rater reliability. The training consisted of the following structured approach. Initially, the significance of clear NCP guidelines and comprehension of desired outcomes was emphasized. Examples of each scale point, 'Not all', 'Some but not all', 'Most but not all', and 'All', were provided to illustrate varying degrees of alignment. In a pilot test of the process, the dietitians engaged in practice sessions where they categorized diverse ChatGPT outputs based on the 4-point Likert scale adopted in this study, elaborating on their reasoning for their evaluations. Open discussion was encouraged to refine their understanding and ensure rating consistency. Feedback and calibration exercises were provided to improve the accuracy of the rating and enhance the understanding of scale nuances. In this study, Statistical Package for Social Sciences 28.0 (SPSS for Windows, 2021, Chicago: SPSS Inc.) was used to calculate the Cohen's weighted kappa to examine the inter-rater (inter-dietitian) reliability of the responses' evaluations. The Cohen's kappa is a form of correlation coefficient which ranges from -1 to $+1$, whereby values ≤ 0 indicate no agreement, 0.01–0.20 as none to slight, 0.21–0.40 as fair, 0.41– 0.60 as moderate, 0.61–0.80 as substantial, and 0.81–1.00 as almost perfect agreement. In this study, the weighted kappa (rather than kappa) was selected given that the rating used to evaluate the responses represented an ordinal scale with more than 2 levels [34].

For the 'NCP' domain, the dietitians examined the chatbot responses according to the following four NCP pillars separately: Nutritional assessment, Diagnosis, Intervention and Monitoring/Evaluation. The 'dietary management' domain was evaluated as per the intervention section of the NCP guidelines. For the 'menu planning' domain, a 1500 Kcal-based diet for an adult woman was used in the prompts. The food items and their corresponding quantities within the menus generated by the Chatbot were entered in the NutritionistPro nutrition analysis program [35]. The menus were analyzed for energy, macro and micronutrients content and were compared to the recommendations outlined in the NCM. In case not specified in the NCM, the Institute of Medicine's Dietary Reference Intakes (DRI) (i.e., Recommended Dietary allowance (RDA) or Adequate Intake (AI) if an RDA is not available) and Acceptable Macronutrient Distribution Ranges (AMDR) references were used for the assessment of adequacy of intakes [36–42].

To examine the quality of the outputs, three main characteristics were evaluated including clarity, coherence and practicality. Similar to accuracy, a four-point Likert scale was used, referring to (1) Poor, (2) Fair, (3) Good and (4) Excellent. A rubric describing the characteristics of the output for each is described in Appendix C.

RESULTS

Table 1 shows the dietitians' average scores for the AI-derived dietary management of T2DM, MetS and its components. Both dietitians

Table 2. Evaluation of the AI-produced Nutrition Care Process for T2DM, the MetS and its components.

	Dietitian1[a]	Dietitian2[a]	Gaps
T2DM			
Nutritional Assessment	1	1	Biochemical data, anthropometrics, medical tests (e.g., assessment of vitamin B12 in patients taking metformin), cardiovascular and renal risk indicators, nutrition-focused physical findings, energy and food intake, diet experience, access to food, medication, and physical activity.
Diagnosis	2	2	PES statement not mentioned or missing nutrition diagnostic terminology for T2DM.
Intervention	2	2	Weight management, CHO counting, glycemic load, specific diet interventions (CHO-modified diet, protein-modified diet, fiber-modified diet, DASH diet, Mediterranean diet), alcohol intake and medication use, adjusting time, type or number of CHO servings.
Monitoring/ Evaluation	1	1	Effectiveness of MNT Plan, glycemic goals for non-pregnant adults with diabetes, laboratory evaluation (e.g., measuring HbA1C at least twice annually).
MetS			
Nutritional Assessment	2	2	Anthropometric measurements (e.g., waist circumference and its cut-offs), weight history, weight goals, food and nutrition-related history, alcohol intake, medications, biochemical data, physical activity.
Diagnosis	3	3	Example PES statements were either not mentioned or incomplete; missing nutrition diagnostic terminology.
Intervention	2	3	Reduction of energy intake, increasing intakes of whole grains and fiber, reducing SFA, TFA and simple sugars, physical activity, relapse-prevention training.
Monitoring/ Evaluation	4	4	None.
Hyperglycemia (or Prediabetes)			
Nutritional Assessment	1	1	Energy and food intake, diet experience, medication, access to food, physical activity, anthropometrics, biochemical data, nutrition-focused physical finding, medical tests and procedures.
Diagnosis	2	2	Example PES statements were either not mentioned or incomplete; didn't include nutrition diagnostic terminology.
Intervention	2	2	Meal planning, reading food labels, physical activity, nutrition counseling, low consumption of red and processed meat, whole-fat dairy products, SFA, TFA, and increased consumption of vegetables, legumes, grains, fruits, and nuts.
Monitoring/ Evaluation	1	1	Effectiveness of MNT plan, laboratory evaluation (e.g., measuring HbA1C at least twice annually).
Obesity			
Nutritional Assessment	2	2	Dietary intake assessment (partial), weight change, nutrition focused physical exams. Biochemical data/tests such as Vitamin D, HbA1c, CRP, and CBC.
Diagnosis	3	3	Example PES statements were either not mentioned or incomplete; missing nutrition diagnostic terminology.
Intervention	2	2	Recommended energy deficit of 500–700 kcal, lean proteins, plant proteins, limiting SFA, sources of omega 3 fatty acids, food supplements, added sugars and refined grains, physical activity, intermittent fasting, approved medication or bariatric surgery.
Monitoring/ Evaluation	3	3	Nutritional status re-assessment (anthropometrics), weight maintenance, monitoring quality of diet and self-monitoring, changes in body composition.
HTN			
Nutritional Assessment	2	3	Missing certain anthropometrics and biochemical data, medications and complementary/ alternative medicine, food and nutrition-related history, nutrition-focused physical findings.
Diagnosis	3	3	Example PES statements were incomplete; missing nutrition diagnostic terminology.
Intervention	3	3	Achieving and maintaining a healthy weight, tips on cutting down sodium, increasing high-potassium foods, physical activity.
Monitoring/ Evaluation	1	1	Monitoring medical condition, medical treatment plans, laboratory values, dietary intake, physical activity, weight and weight changes.
Hypertriglyceridemia			
Nutritional Assessment	1	2	Energy intake, protein intake, fat (high fat foods, TFA and omega 3 sources), physical activity, specific biochemical data (e.g., vitamin D, HbA1c, Liver enzymes), anthropometric measurements, nutrition-focused physical exam, complementary/alternative medicine.
Diagnosis	3	3	Example PES statements were either not mentioned or incomplete; missing nutrition diagnostic terminology.
Intervention	1	1	Fat intake: 25–35% of energy intake, increasing soluble fiber, adding plant stanols/sterols, supplementing with fish oil, alcohol intake, limiting SFA to less than 7% of energy intake, choosing MUFAs and PUFAs, encouraging whole grains and high-fiber foods, examples of foods high in cholesterol, limiting cholesterol to < 200 mg, physical activity.

Table 2. continued

	Dietitian1[a]	Dietitian2[a]	Gaps
Monitoring/ Evaluation	2	2	Nutritional reassessment (changes in weight, percentage of body fat, and waist circumference), monitoring medical condition, medical treatment plans, current laboratory values (serum TG, high-density lipoprotein cholesterol level, and fasting blood sugar).
Low HDL levels			
Nutritional Assessment	2	2	Dietary intake assessment, use of dietary supplements, medications and complementary/ alternative medicine, nutrition focused physical findings, physical activity, anthropometrics, some biomedical data and tests (e.g electrolytes, inflammation indicators).
Diagnosis	3	3	PES statement was either not mentioned or incomplete (missing diagnosis terms).
Intervention	2	2	Recommended % weight loss, dietary patterns (DASH diet/ Mediterranean diet), physical activity, and the inclusion of viscous fiber, plants stanols/sterols, and fish oil supplements.
Monitoring/ Evaluation	2	2	Reassessment parameters (intakes, biochemical data, physical activity, body composition/growth/ weight history).

AI artificial intelligence, *CBC* complete blood count, *CHO* carbohydrates, *CRP* c-reactive protein, *DASH* Dietary Approaches to Stop HTN, *HDL* high-density lipoprotein, *HTN* hypertension, *MetS* metabolic syndrome, *MNT* medical nutrition therapy, *MUFA* mono unsaturated fatty acids, *PES* problem, etiology, signs and symptoms, *PUFA* poly unsaturated fatty acids, *SFA* saturated fatty acids, *T2DM* type 2 diabetes mellitus, *TG* triglycerides, *TFA* trans fatty acids.
[a]1= none; 2= some but not all; 3= most but not all; 4= all.

provided identical ratings for obesity, hypertriglyceridemia, MetS, hyperglycemia (or prediabetes), HTN, and low HDL levels while one-point differences were noted for T2DM. The Cohen's weighted kappa for inter-rater evaluation was 0.59, $p = 0.010$.

The gaps in the AI-derived dietary management outputs for T2DM, MetS and its components, as identified by the dietitians are summarized in Table 1. Common gaps for hyperglycemia (or prediabetes) and T2DM were related to providing appropriate amounts/servings from varied food groups, planning and distribution of meals. For T2DM, additional gaps were identified specific to weight loss management, macronutrient intakes, specific diet interventions (such as CHO and fiber-modified diets, protein-modified diets, the Mediterranean diet, and the DASH diet [Dietary Approaches to Stop Hypertension]), CHO counting, glycaemic load, and hypoglycaemia-related considerations/ management. An important gap observed in the outputs for obesity, MetS, and hypertriglyceridemia levels was related to energy intake, as these recommendations failed to address energy intake modification as a target in the dietary management of these conditions. For low HDL-levels, an important gap was related to weight loss recommendations. The AI outputs failed to address specific critical areas in the dietary management of the cardiometabolic conditions considered in this study such as increasing the intake of nutrient-dense foods, lean and plant proteins, and omega 3 fatty acids; decreasing the consumption of SFA, added sugars, and refined grains. Intermittent fasting was not tackled, despite it being addressed by the NCM as a potential intervention for obesity. In addition, the AI recommendations for HTN did not provide strategies to cut down on sodium intakes to reduce blood pressure. Similarly, the outputs missed recommendations pertinent to the intakes and sources of specific types of fat such as omega-3 fatty acids, fish and fish oils as well as plant stanols/sterols in the case of high triglycerides (TG) and low HDL (Table 1).

Table 2 shows the evaluation of the AI-produced NCP for T2DM, MetS and its components. In this table, for each of the conditions, the scores assigned by the dietitians for the Nutritional assessment, Diagnosis, Intervention and Monitoring/Evaluation are presented. Each score represents the average of the three prompts which were used in the study. Overall, dietitian 1 and dietitian 2 had similar ratings for almost all the metabolic disorders. Even when discrepant (such is the case for nutrition assessment recommendations for HTN and high TG), the difference in the scores did not exceed 1 point.

With regards to the Nutritional Assessment part of the NCP, for most conditions, anthropometric assessment and their corresponding cut-offs were missing. In addition, the biochemical assessment part of the Nutritional Assessment did not include important indicators such as vitamin B12 for T2DM patients taking metformin. Another important gap in the assessment was the evaluation of the patient history including the use of complementary and alternative medications for HTN, high TG and low HDL. Furthermore, the outputs did not address dietary assessment, especially energy, protein intake, specific types of fat such as trans-fat and omega-3 fatty acids (in the case of high TG), alcohol intake (for the MetS) and use of dietary supplements (for low HDL). The Cohen's weighted kappa for the agreement between the two dietitians in this section was 0.533, $p = 0.05$

For the Diagnosis section of the NCP, although the AI generated a diagnosis statement, it was deemed to be an incomplete problem, etiology, signs and symptoms (PES) statement with missing key diagnostic terms as observed across all metabolic disorders considered. For example, in the case of T2DM, the AI-generated PES statement was 'impaired glucose tolerance related to inconsistent CHO intake'. While this statement briefly mentioned the Problem (impaired glucose tolerance) and the Etiology (inconsistent CHO intake), it made no reference to the Signs and Symptoms, as such it was deemed incomplete. In this section, both dietitians gave similar evaluations, as such the Cohen's weighted kappa was 1.00 with a *p*-value of 0.008.

Regarding the Intervention section of the NCP, similar gaps to those described in Table 1 were identified. For instance, for T2DM, the AI-generated output did not address weight management, CHO counting or type/number of CHO servings and glycemic load, all of which are hallmarks of dietary interventions in T2DM. Alarmingly, the AI missed the reduction in energy intake in both MetS and obesity. A common gap is related to fiber intake, whereby for all conditions considered, the AI did not address the need to increase fiber or consuming whole grain products. In addition to dietary intake, the AI output consistently missed important elements of the intervention including alcohol intake and physical activity. Similar to the Diagnosis section of the NCP, in this section the Cohen's weighted kappa was 1.00 with a *p*-value of 0.008; since the ratings of both dietitians were identical.

Except for the MetS, important gaps were identified in the AI-generated recommendations for the Monitoring and Evaluation part of the NCP. For instance, despite high scores for obesity (3 for both dietitians), the fact that the AI did not indicate the need for nutritional status reassessment following the intervention is critical. In the case of hypertriglyceridemia and low HDL levels, the score for this section was 2 for both dietitians, with neither nutritional reassessment nor biochemical assessment being

SPRINGER NATURE

mentioned. For the remaining conditions, the scoring for this section was 1, reflecting a 'not at all' grade; whereby none of the monitoring and evaluation principles were addressed by the AI-generated outputs.

Table 3 displays the energy and selected macro- and micronutrients content of AI-derived menus designed for patients with T2DM, MetS or its individual components. Concerning energy, there existed important differences between the AI-generated diets and the specified 1500 kcal menu plan, such as +361 Kcal in the case of hypertriglyceridemia. Except for T2DM, the diets were lower than the recommendations for CHO, reaching as low as 36.8% of energy intake in the case of hypertriglyceridemia. Fat, on the other hand, was higher than the recommended amounts for all conditions, except for obesity where it was within range. Similarly, cholesterol exceeded the 200 mg recommendations for T2DM, hyperglycemia (or prediabetes), MetS, obesity, and hypertriglyceridemia. The AI-derived menus met the requirements for most of the micronutrients except for calcium and vitamin D, where the diets were falling below recommendations. (Table 3).

Regarding the quality of the outputs, for all the 63 outputs considered in this study, the clarity, coherence and practicality were rated between 3 and 4, indicating Good to Excellent; with none of the output rated as Poor or Fair.

DISCUSSION

This study showed that the AI-generated outputs for the dietary management of T2DM and the MetS were often incomplete or discordant with the NCM recommendations, presenting major gaps of significant implications in patient care and highlighting the need to further improve the accuracy of these output. Similar observations have been reported by previous studies in non-nutrition related medical fields. For instance, Chen et al. (2023) evaluated a large language models chatbot's therapeutic recommendations for breast, prostate, and lung cancer, and their concordance with the National Comprehensive Cancer Network (NCCN) guidelines. The authors found that the chatbot's recommendations were, at least partially, non-aligned with the NCCN guidelines and the output included "hallucinated" responses [19]. Similarly, in a recent study where endocrinologists asked questions to ChatGPT focusing on assessment and treatment options for obesity in T2DM, ChatGPT output was found to be compatible with the guidelines for obesity assessment while its compatibility was lower in sections pertinent to nutritional, medical, and surgical therapeutic strategies [12]. In the nutrition field, Qarajeh et al. (2023) used four different AI models to investigate the efficacy of different AI models in correctly identifying the potassium and phosphorus content of foods for chronic kidney disease patients. The AI-derived results were compared to the Mayo Clinic Renal Diet Handbook's recommendations [23]. While being of promising potential, the results suggested variations among the different AI models' outputs and a diversity in their range of accuracy, thus emphasizing the need for human oversight when using AI in the field of nutritional education and recommendations [23]. Furthermore, Aiumtrakul et al. (2024) examined the reliability of chatbots in classifying foods according to their oxalate content [24]. The study findings highlighted significant variations in the accuracy of AI-models in classifying dietary oxalate content and emphasized the need for further improvements in chatbot algorithms for therapeutic accuracy [24].

Regarding dietary interventions, the observed gaps in our study were relatively similar when using the "dietary management" or the "NCP" domains. For example, weight loss recommendations along with guidance on achieving energy deficit were generally missing or incomplete in the AI-generated outputs. More specifically, the AI-derived dietary management outputs for obesity, MetS, and hypertriglyceridemia levels did not address energy intake modification, and for low HDL-levels, an important gap was related to weight loss recommendations. This is concerning given the well-established pathophysiologic links between excess weight, insulin resistance, MetS and T2DM [43, 44]. Indeed, weight loss is one of the most impactful therapeutic interventions for cardiometabolic conditions. The 2023 Standards of Care in Diabetes indicate that relatively modest weight loss (3–7% of baseline body weight), significantly improves glycemic control and cardiometabolic status, while larger losses (exceeding 10% body weight) possibly lead to disease remission [8]. Effective strategies for weight loss involve the establishment of an energy deficit (500–1000 kcal/day), via hypocaloric diets and increased physical activity [8, 45]. Notably, AI-derived outputs were lacking appropriate physical activity recommendations for all the cardiometabolic conditions studied in this paper, except for T2DM. Hence, despite ChatGPT's ability to generate 'contextually relevant responses' due to its extensive training on large datasets, it still encounters challenges in finding interdisciplinary connections and providing holistic lifestyle interventions that experienced dietitians often provide as part of patient counseling.

Dietary intervention information generated via ChatGPT were also often incomplete in terms of guidance on specific nutrients of concern such as sodium intake in the case of HTN, the amount, distribution and quality of CHO for T2DM, and the intake of SFA and added sugar for obesity [7, 8]. Notably, information pertinent to CHO counting were not included in the AI-derived dietary management output although CHO counting is a cornerstone of therapeutic diets in cases of T2DM [46]. In fact, CHO counting, even in its most basic form, has been consistently linked to better glycemic control among patients with T2DM and hence better health outcomes [47, 48]. Furthermore, recommendations related to increasing the intakes of other nutrients known to improve cardiometabolic health was also incomplete/missing such as those pertinent to dietary fiber, [7, 8]. For example, the AI-derived output on dietary intervention within the NCP did not address the need to increase fiber and soluble/viscous fiber or to consume whole grain products for all the considered conditions. The promotion of dietary fiber intake is a crucial component of the management of cardiometabolic abnormalities since, through its colonic and hormonal effects, adequate intake of dietary fiber can increase insulin sensitivity, enhance fat oxidation, and decrease cardiometabolic risk [49]. Similarly, recommendations pertinent to omega-3 fatty acids from fatty fish were missing or incomplete for most of the studied cardiometabolic health conditions, despite their inclusion in the NCM and in clinical recommendations issued by scientific bodies [7, 8]. The AI-derived output was also lacking recommendations addressing dietary intake as a whole, in the form of dietary patterns. This is in contrast to the NCM recommendations as well as recent recommendations published by the ADA [8] and the AHA [50]. For instance, the ADA made reference to specific dietary interventions such as the Mediterranean or the DASH diet as potential dietary patterns that can be adopted by T2DM patients to foster weight loss and improve cardiometabolic profile [8]. A scientific statement from the AHA also showed that several dietary patterns strongly align with the 2021 AHA Dietary Guidance [50]. Interventions based on dietary patterns are gaining increasing popularity as the guidance provided by this approach is often clearer and easier to follow by the patient as compared to recommendations based on nutrients or individual foods [51].

In the present study, AI-generated outputs for nutritional assessment within the NCP were incomplete for most of the considered conditions. More specifically, the AI-derived output was lacking in terms of dietary assessment, such as the intakes of energy, protein and specific types of fat (in the case of high TG), alcohol intake (for the MetS) and use of dietary supplements (for low HDL). Anthropometric assessment in terms of indicators and cutoffs were also missing from the NCP for T2DM, MetS,

Table 3. Evaluation of energy and selected macro- and micronutrients contents of AI derived menus designed for the management of T2DM, the MetS or its individual components.

	T2DM		MetS		Hyperglycemia (or Prediabetes)		Obesity		HTN		High TG		Low HDL	
	Chat-GPT	Recommendations	Chat-GPT	Recommendations	Chat-GPT	Recommendations	Chat-GPT	Recommendations	Chat-GPT	Recommendations	Chat-GPT	Recommendations	Chat-GPT	Recommendations
Energy (kcal)	1462	1500	1679.3	1500	1627	1500	1581	1500	1528.5	1500	1861.4	1500	1698.4	1500
CHO[a]	35.57	26–45	38.14	45–65	37.2	45–65	42.48	45–65	42.58	45–65	36.8	45–65	40	45–65
Protein[a]	27.91	15–20	28.75	10–35	30.0	10–35	29.62	10–35	25.99	10–35	30.6	15–20	21.8	10-35
Fat[a]	39.91	20–35	36.01	20–35	35.4	20–35	31.4	20–35	35.23	20–35	35.9	25–35	43.2	20–35
SFA[a]	7.15	7	6.88	7	5.96	7	5.88	7	5.75	7	6.58	7	6.32	7%
TFA[a]	0.07	<0.01	0.06	<0.01	0.06	<0.01	0.04	<0.01	0.05	<0.01	0.05	<0.01	0.06	<0.01
Cholesterol (mg)	283.4	200	484.6	200	266.4	200	295.9	200	143.4	200	456.4	200	103.7	200
Fiber (g)	30.28	21–25	40.83	21–25	35.1	21–25	39.81	21–25	40.09	21–25	41.2	21–25	43.5	21–25
Calcium (mg)	673.4	1000–1200	744.6	1000–1200	811.3	1000–1200	917.3	1000–1200	800	1000–1200	1047.6	1000–1200	767	1000–1200
Iron (mg)	13.38	8–18	18.47	8–18	15.3	8–18	19	8–18	18.48	8–18	20.9	8–18	13.4	8–18
Zinc (mg)	9.04	8	11.81	8	11.1	8	11.41	8	10.42	8	12.9	8	10.1	8
Sodium (g)	0.71	1.2–1.5	0.82	1.2–1.5	0.78	1.2–1.5	0.91	1.2–1.5	0.82	1.2–1.5	1.43	1.2-1.5	0.9	1.2-1.5
Vitamin A (μg)	2008.7	700	2273.9	700	1637.3	700	2350.2	700	2320.9	700	3084.1	700	2221	700
Vitamin C (mg)	203.7	75	205.3	75	220.4	75	223.8	75	197.9	75	258.3	75	293.1	75
Vitamin D (μg)	0.81	15–20	1.95	15–20	1.36	15–20	1.82	15–20	0.09	15–20	1.59	15–20	0.03	15–20
Vitamin E (mg)	15.07	15	15.87	15	18	15	17.4	15	15.81	15	19.4	15	21.7	15
Folate (μg)	659.2	400	946.69	400	730.9	400	937.34	400	811.03	400	1056.72	400	759.49	400
Vitamin B12 (μg)	4.71	2.4	4.79	2.4	5.27	2.4	4.63	2.4	3.68	2.4	6.43	2.4	3.93	2.4

Recommendations were based on the NCM [20] or the Institute of Medicine's DRI and AMDR [22–28].
AI artificial intelligence, *CHO* carbohydrates, *HDL* high-density lipoprotein, *HTN* hypertension, *MetS* metabolic syndrome, *SFA* saturated fatty acids, *T2DM* type 2 diabetes mellitus, *TG* triglycerides, *TFA* trans fatty acids.
[a]expressed as % of total energy intake.

hypertriglyceridemia and the monitoring/evaluation component for obesity. Acknowledging the crucial role that nutritional assessment plays in identifying nutrition-related problems and their causes, as well as developing tailored dietary interventions, the observed gaps in AI-derived outputs call for caution in relying solely on ChatGPT for the assessment of patient's nutritional status and the development of nutritional care plans [52, 53].

There were also gaps in the AI-derived Nutrition diagnosis outputs, with the PES statements not being mentioned or having missing diagnostic terminology. Accurate nutrition diagnosis and clearly stated PES statements represent a crucial component of the patient care plan, allowing for the determination of subsequent steps within the NCP and the development of tailored nutrition interventions [54, 55]. More specifically, accurate nutrition diagnosis allows to identify and prioritize problems that can most likely be resolved or improved by a nutrition intervention, evaluate if the "root cause" can be addressed by the intervention and select the specific signs and symptoms from the assessment data to allow for the monitoring and tracking of whether the problem was resolved or improved [56]. Quality improvement literature shows that incomplete nutrition diagnoses and the lack of standardized approaches in the identification of nutrition problems increase the variation and unpredictability of therapeutic outcomes and hence the effectiveness of patient care [56]. The AI-derived output had also significant gaps in the monitoring and evaluation step, a crucial step within the NCP that assesses patients' progress and serves as a basis for adjustments at the level of previous diagnoses or nutrition interventions [57].

In this study, ChatGPT was prompted to provide a 1500 kcal 1-day menu for different diets (conditions). In alignment with dietary recommendations, and as observed by previous studies [58], the generated menus contained non-starchy vegetables, lean protein foods, and CHO foods, but upon nutritional analysis, these menus showed that the energy amounts sometimes diverged from the assigned 1500 kcal (e.g., by + 361 Kcal in the case of hypertriglyceridemia), and the amounts of CHO and fat were often discordant with the dietary recommendations. The observed low amount of CHO in some of the menus is in line with findings reported by Chatelan, et al., 2023 [58] where a ChatGPT-derived 1-day diet menu for T2DM was described "to be inspired by ketogenic diets". In addition, in our study, the dietary fat amounts were high, reaching 43% of energy intake in the menu for low HDL management and 40% for T2DM. Such high levels of fat intake may potentially exacerbate cardiometabolic abnormalities including abdominal obesity, blood pressure and dyslipidemia [59]. As for micronutrients, certain discrepancies were noted between AI-derived outputs and nutritional recommendations particularly for calcium and vitamin D intakes, whereby the ChatGPT-generated menus included very low levels of vitamin D and inadequate calcium intake. This is concerning, given the high prevalence of micronutrient deficiencies in the general population worldwide, and among patients with T2DM and the MetS [60].

With regards to clarity/quality of the AI-derived outputs, the findings of this study showed that the outputs of ChatGPT were of good to excellent clarity, coherence and practicality. In fact, by design, ChatGPT has been trained on extensive human language datasets, and several studies confirmed its ability to produce clear and coherent text output [61–63].

The present study suggests few recommendations to be considered when using ChatGPT for the dietary management of T2DM and the MetS. Although ChatGPT is an increasingly available resource, enhancing accessibility to nutrition information online, evaluating the comprehensiveness and accuracy of the responses remains crucial for AI chatbot users [58]. In our study, although there were minor disagreements between the dietitians' scores, these scores were not always identical, highlighting the potential challenges of interpreting the AI-generated therapeutic strategies. The fact that ChatGPT-derived responses lack references further limits the possibility to ascertain the credibility of the information [58]. Investigations evaluating the gaps of ChatGPT, such as the current study, are crucial to raise awareness among dietitians regarding AI-generated information and to caution healthcare providers from relying solely on ChatGPT for dietetic advice [58]. Dietitians and health care practitioners also have the responsibility of educating their clients/patients about the risks of the use of this technology for disease treatment or prevention [64] as the use of AI-chatbots could propagate misinformation due to its widespread use for self-education and self-management by patients and the general public [19]. In an opinion paper, Arslan (2024) noted that although chatbots can be harnessed for the provision of real-time nutritional information, a cautious and critical approach should be adopted in its utilization [22]. A real-world example brought forward by Arslan is an application where ChatGPT was integrated within online nutritional counseling platforms as an initial advisory tool [25], allowing users to obtain initial guidance and foundational knowledge prior to consulting with human experts (i.e., dietitians). There are also other opportunities for ChatGPT to assist nutritionists and dietitians. For instance, AI models can assist in providing a quick, 24/7 second opinion, when utilized advisedly and when the dietitian is able to correctly define an issue, ask pertinent questions, refine the prompts and gauge the accuracy of the output. It can also help in brainstorming certain ideas (such as nutrition education objectives, research hypotheses, etc.), providing quick summaries of texts, and drafting texts that can be easily understood by the patients and hence can be used for effective nutritional counseling [58]. In addition, given that ChatGPT and other similar AI systems are trained using large sets of training data and typically provide responses based on the data they were trained on, it is recommended to add more evidence-based nutritional knowledge and recommendations to the training data of AI and to better synchronize AI algorithms with nutrition guidelines [22].

The findings of this study should be interpreted in light of the following limitations. First, AI-generated answers to the same prompt may vary over time due to the regular optimizations of ChatGPT [12, 58, 65], making it difficult to know which exact information a health professional/patient would receive, especially if additional data are provided related to sex, age, food preferences, and medical history [58]. The continuously changing nature of the ChatGPT presents a challenge in assessing the reliability of the information, especially in studies that adopt a zero-shot approach [12, 66]. Second, differences in prompt engineering like wording, design and implementation could generate different results. In this study, only 3 versions of the question were designed. Increasing the variety and complexity of the prompt could provide a more comprehensive evaluation of the ChatGPT performance [67]. Third, including a human comparison group could have provided a more contextual appraisal of the ChatGPT. That said, this study was based on the premise that dietitians, especially registered dietitians, as well as the ChatGPT are expected to strictly adhere to the NCP guidelines in the management of diseases. Finally, in this study, ChatGPT outputs were evaluated in a controlled setting. Future studies can further assess ChatGPT's performance in real-world clinical scenarios and actual patients' cases.

In conclusion, despite their clarity, the ChatGPT-derived outputs in the field of T2DM and the MetS nutritional management presented several gaps, including weight loss recommendations, energy deficit, anthropometric assessment, specific nutrients of concern, the adoption of specific dietary interventions, physical activity recommendations, diagnostic documentation statements, and monitoring and evaluation. In the 1500 kcal one-day menus, the amounts of CHO, fat, vitamin D and calcium were discordant with dietary recommendations. These findings suggest that ChatGPT, and potentially other future AI chatbots, react to the user's prompts in "a human-like" way, but cannot replace the dietitians' expertise and critical judgment. While healthcare practitioners may consult this increasingly available technology for various purposes, they must also be cautious about relying solely on AI chatbots in clinical practice and should collectively raise awareness about associated risks.

DATA AVAILABILITY

Data will be made available by the corresponding author on reasonable request.

REFERENCES

1. World Health Organization. Noncommunicable diseases: key facts. 2023. https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases#:~:text=Noncommunicable%20diseases%20(NCDs)%20kill%2041,%2D%20and%20middle%2Dincome%20countries. Accessed 12 Dec 2023.
2. World Health Organization. Diabetes: key facts. 2023. https://www.who.int/news-room/fact-sheets/detail/diabetes. Accessed 12 Dec 2023.
3. Ong KL, Stafford LK, McLaughlin SA, Boyko EJ, Vollset SE, Smith AE. et al. Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: a systematic analysis for the Global Burden of Disease Study 2021. Lancet. 2023;402:203–34. https://doi.org/10.1016/S0140-6736(23)01301-6.
4. Fahed G, Aoun L, Bou Zerdan M, Allam S, Bou Zerdan M, Bouferraa Y, et al. Metabolic syndrome: updates on pathophysiology and management in 2021. Int J Mol Sci. 2022;23:786. https://doi.org/10.3390/ijms23020786.
5. Wilson PW, D'Agostino RB, Parise H, Sullivan L, Meigs JB. Metabolic syndrome as a precursor of cardiovascular disease and type 2 diabetes mellitus. Circulation. 2005;112:3066–72. https://doi.org/10.1161/CIRCULATIONAHA.105.539528.
6. Noubiap JJ, Nansseu JR, Lontchi-Yimagou E, Nkeck JR, Nyaga UF, Ngouo AT, et al. Geographic distribution of metabolic syndrome and its components in the general adult population: A meta-analysis of global data from 28 million individuals. Diabetes Res Clin Pract. 2022;188:109924. https://doi.org/10.1016/j.diabres.2022.109924.
7. Grundy SM, Hansen B, Smith SC Jr, Cleeman JI, Kahn RA, Participants C. Clinical management of metabolic syndrome: report of the American Heart Association/National Heart, Lung, and Blood Institute/American Diabetes Association conference on scientific issues related to management. Circulation. 2004;109:551–56. https://doi.org/10.1161/01.CIR.0000112379.88385.67.
8. American Diabetes Association. Standards of care in diabetes—2023. Diabetes Care 2023;46. https://doi.org/10.2337/dc23-Sint.
9. Guzmán A, Navarro E, Obando L, Pacheco J, Quirós K, Vásquez L, et al. Effectiveness of interventions for the reversal of a metabolic syndrome diagnosis: an update of a meta-analysis of mixed treatment comparison studies. Biomedica. 2019;39:647–62. https://doi.org/10.7705/biomedica.4684.
10. Thirunavukarasu AJ, Ting DSJ, Elangovan K, Gutierrez L, Tan TF, Ting DSW. Large language models in medicine. Nat Med. 2023;29:1930–40. https://doi.org/10.1038/s41591-023-02448-8.
11. Else H. Abstracts written by ChatGPT fool scientists. Nature. 2023;613:423. https://doi.org/10.1038/d41586-023-00056-7.
12. Barlas T, Altinova AE, Akturk M, Toruner FB. Credibility of ChatGPT in the assessment of obesity in type 2 diabetes according to the guidelines. Int J ObeS. 2024;48:271–75. https://doi.org/10.1038/s41366-023-01410-5.
13. Sivasubramanian J, Hussain SMS, Muthuprakash SV, Periadurai ND, Mohanram K, Surapaneni KM. Analysing the clinical knowledge of ChatGPT in medical microbiology in the undergraduate medical examination. Indian J Med Microbiol. 2023;45:100380. https://doi.org/10.1016/j.ijmmb.2023.100380.
14. Seney V, Desroches ML, Schuler MS. Using ChatGPT to teach enhanced clinical judgment in nursing education. Nurse Educ. 2023;48:124. https://doi.org/10.1097/NNE.0000000000001383.
15. Huh S. Are ChatGPT's knowledge and interpretation ability comparable to those of medical students in Korea for taking a parasitology examination?: a descriptive study. J Educ Eval Health Prof. 2023;20. https://doi.org/10.3352/jeehp.2023.20.1.
16. Bhayana R, Krishna S, Bleakney RR. Performance of ChatGPT on a radiology board-style examination: Insights into current strengths and limitations. Radiology. 2023;307:230582. https://doi.org/10.1148/radiol.230582.
17. Sedaghat S. Success through simplicity: what other artificial intelligence applications in medicine should learn from history and ChatGPT. Ann Biomed Eng. 2023;51:2657–58. https://doi.org/10.1007/s10439-023-03287-x.
18. Arslan S. Exploring the potential of chat GPT in personalized obesity treatment. Ann Biomed Eng. 2023;51:1887–88. https://doi.org/10.1007/s10439-023-03227-9.
19. Chen S, Kann BH, Foote MB, Aerts HJWL, Savova GK, Mak RH, et al. Use of artificial intelligence chatbots for cancer treatment information. JAMA Oncol. 2023;9:1459–62. https://doi.org/10.1001/jamaoncol.2023.2954.
20. Haupt CE, Marks M. AI-generated medical advice—GPT and beyond. JAMA. 2023;329:1349–50. https://doi.org/10.1001/jama.2023.5321.
21. Au Yeung J, Kraljevic Z, Luintel A, Balston A, Idowu E, Dobson RJ, et al. AI chatbots not yet ready for clinical use. Front Digit Health. 2023;5. https://doi.org/10.3389/fdgth.2023.1161098.
22. Arslan S. Decoding dietary myths: the role of ChatGPT in modern nutrition. Clin Nutr ESPEN. 2024;60:285–88. https://doi.org/10.1016/j.clnesp.2024.02.022.
23. Qarajeh A, Tangpanithandee S, Thongprayoon C, Suppadungsuk S, Krisanapan P, Aiumtrakul N, et al. AI-powered renal diet support: performance of ChatGPT, Bard AI, and Bing Chat. Clin Pract. 2023;13:1160–72. https://doi.org/10.3390/clinpract13050104.
24. Aiumtrakul N, Thongprayoon C, Arayangkool C, Vo KB, Wannaphut C, Suppadungsuk S, et al. Personalized Medicine in urolithiasis: AI chatbot-assisted dietary management of oxalate for kidney stone prevention. J Pers Med. 2024;14:107. https://doi.org/10.3390/jpm14010107.
25. Javaid M, Haleem A, Singh RP. ChatGPT for healthcare services: an emerging stage for an innovative perspective. TBench. 2023;3:100105. https://doi.org/10.1016/j.tbench.2023.100105.
26. Firat M. What ChatGPT means for universities: perceptions of scholars and students. J Appl Learn Teach. 2023;6:57–63. https://doi.org/10.37074/jalt.2023.6.1.22.
27. Bahrini A, Khamoshifar M, Abbasimehr H, Riggs RJ, Esmaeili M, Majdabadkohne RM, et al. editors. ChatGPT: Applications, opportunities, and threats. 2023 Systems and Information Engineering Design Symposium (SIEDS). IEEE; 2023. https://doi.org/10.1109/SIEDS58326.2023.10137850.
28. Morita PP, Abhari S, Kaur J, Lotto M, Miranda PADSES, Oetomo A. Applying ChatGPT in public health: a SWOT and PESTLE analysis. Front Public Health 2023;11:1225861. https://doi.org/10.3389/fpubh.2023.1225861.
29. Garcia MB. ChatGPT as a virtual dietitian: exploring its potential as a tool for improving nutrition knowledge. Appl Syst Innov. 2023;6:96 https://doi.org/10.3390/asi6050096.
30. Ray PP. ChatGPT: a comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. Internet Things Cyber-Phys Syst. 2023;3:121–54. https://doi.org/10.1016/j.iotcps.2023.04.003.
31. Academy of Nutrition and Dietetics. Nutrition Care Manual. https://www.nutritioncaremanual.org/welcome?_lid=A9BE98DA-E4F5-7840-E17A623293D73F35. Accessed 2 Oct 2023.
32. Johnson D, Goodman R, Patrinely J, Stone C, Zimmerman E, Donald R, et al. Assessing the accuracy and reliability of AI-generated medical responses: an evaluation of the Chat-GPT model [Preprint]. Res Sq. 2023;rs.3:rs-2566942. https://doi.org/10.21203/rs.3.rs-2566942/v1.
33. University of North Florida Digital Pressbooks. Evaluating ChatGPT-Generated Content; 2023.
34. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med. 2012;22:276–82. https://doi.org/10.11613/BM.2012.031.
35. AXXYA. NutritionistPro software; https://nutritionistpro.com/.
36. Institute of Medicine. Dietary Reference Intakes for Calcium, Phosphorus, Magnesium, Vitamin D, and Fluoride; The National Academies Press: Washington DC, US. 1997. https://pubmed.ncbi.nlm.nih.gov/23115811/.
37. Institute of Medicine. Dietary Reference Intakes for Thiamin, Riboflavin, Niacin, Vitamin B6, Folate, Vitamin B12, Pantothenic Acid, Biotin, and Choline. Washington DC, US: The National Academies Press; 1998. https://www.ncbi.nlm.nih.gov/books/NBK114310/.
38. Institute of Medicine. Dietary Reference Intakes for Vitamin C, Vitamin E, Selenium, and Carotenoids; The National Academies Press: Washington DC, US. 2000. https://pubmed.ncbi.nlm.nih.gov/25077263/.
39. Institute of Medicine. Dietary Reference Intakes for Vitamin A, Vitamin K, Arsenic, Boron, Chromium, Copper, Iodine, Iron, Manganese, Molybdenum, Nickel, Silicon, Vanadium, and Zinc; The National Academies Press: Washington DC, US. 2001. https://www.ncbi.nlm.nih.gov/books/NBK222310/.
40. Institute of Medicine. Dietary Reference Intakes for Water, Potassium, Sodium, Chloride, and Sulfate; The National Academies Press: Washington DC, US. 2005. https://nap.nationalacademies.org/catalog/10925/dietary-reference-intakes-for-water-potassium-sodium-chloride-and-sulfate.
41. Institute of Medicine. Dietary Reference Intakes for Calcium and Vitamin D; The National Academies: Washington DC, US. 2011. https://www.ncbi.nlm.nih.gov/books/NBK56070/.
42. Institute of Medicine. Dietary Reference Intakes for Energy, Carbohydrate. Fiber, Fat, Fatty Acids, Cholesterol, Protein, and Amino Acids. Washington DC, US: The National Academies Press; 2005. https://doi.org/10.17226/10490.
43. Jin X, Qiu T, Li L, Yu R, Chen X, Li C, et al. Pathophysiology of obesity and its associated diseases. Acta Pharm Sin B. 2023;13:2403–24. https://doi.org/10.1016/j.apsb.2023.01.012.
44. Ruze R, Liu T, Zou X, Song J, Chen Y, Xu R, et al. Obesity and type 2 diabetes mellitus: connections in epidemiology, pathogenesis, and treatments. Front Endocrinol. 2023;14:1161521. https://doi.org/10.3389/fendo.2023.1161521.

45. Eckel RH, Alberti KG, Grundy SM, Zimmet PZ. The metabolic syndrome. Lancet. 2010;375:181–83. https://doi.org/10.1016/S0140-6736(09)61794-3.
46. American Diabetes Association. Standards of care in diabetes. Diabetes Care 2023;46. https://diabetesjournals.org/care/issue/46/Supplement_1.
47. Bowen ME, Cavanaugh KL, Wolff K, Davis D, Gregory RP, Shintani A, et al. The diabetes nutrition education study randomized controlled trial: a comparative effectiveness study of approaches to nutrition in diabetes self-management education. Patient Educ Couns. 2016;99:1368–76. https://doi.org/10.1016/j.pec.2016.03.017.
48. Martins MR, Ambrosio ACT, Nery M, de Cássia Aquino R, Queiroz MS. Assessment guidance of carbohydrate counting method in patients with type 2 diabetes mellitus. Prim Care Diabetes. 2014;8:39–42. https://doi.org/10.1016/j.pcd.2013.04.009.
49. Anderson JW, Baird P, Davis Jr RH, Ferreri S, Knudtson M, Koraym A, et al. Health benefits of dietary fiber. Nutr Rev. 2009;67:188–205. https://doi.org/10.1111/j.1753-4887.2009.00189.x.
50. Gardner CD, Vadiveloo MK, Petersen KS, Anderson CA, Springfield S, Van Horn L, et al. Popular dietary patterns: alignment with American Heart Association 2021 dietary guidance: a scientific statement from the American Heart Association. Circulation. 2023;147:1715–30. https://doi.org/10.1161/CIR.0000000000001146.
51. Slattery ML. Defining dietary consumption: is the sum greater than its parts? Am J Clin Nutr. 2008;88:14–15. https://doi.org/10.1093/ajcn/88.1.14.
52. Younis HA, Eisa TAE, Nasser M, Sahib TM, Noor AA, Alyasiri OM, et al. A systematic review and meta-analysis of artificial intelligence tools in medicine and healthcare: applications, considerations, limitations, motivation and challenges. Diagnostics. 2024;14:109. https://doi.org/10.3390/diagnostics14010109.
53. Mu Y, He D. The potential applications and challenges of ChatGPT in the medical field. Int J Gen Med. 2024;17:817–26. https://doi.org/10.2147/IJGM.S456659.
54. Colin C, Arikawa A, Lewis S, Cooper M, Lamers-Johnson E, Wright L, et al. Documentation of the evidence-diagnosis link predicts nutrition diagnosis resolution in the Academy of Nutrition and Dietetics' diabetes mellitus registry study: A secondary analysis of Nutrition Care Process outcomes. Front Nutr. 2023;10:1011958. https://doi.org/10.3389/fnut.2023.1011958.
55. Skipper A. Applying the nutrition care process: nutrition diagnosis and intervention. Support Line. 2007;29:12–23.
56. American Dietetic Association. Nutrition Diagnosis: A Critical Step in the Nutrition Care Process 2006. https://www.andeal.org/files/File/Nutrition%20Diagnosis.pdf.
57. Writing Group of the Nutrition Care Process/Standardized Language Committee. Nutrition care process and model part I: the 2008 update. J Am Diet Assoc. 2008;108:1113–17. https://doi.org/10.1016/j.jada.2008.04.027.
58. Chatelan A, Clerc A, Fonta P-A. ChatGPT and future artificial intelligence Chatbots: what may be the influence on credentialed nutrition and dietetics practitioners? J Acad Nutr Diet. 2023;123:1525–31. https://doi.org/10.1016/j.jand.2023.08.001.
59. Powell-Wiley TM, Poirier P, Burke LE, Després J-P, Gordon-Larsen P, Lavie CJ, et al. Obesity and cardiovascular disease: a scientific statement from the American Heart Association. Circulation. 2021;143:e984–e1010. https://doi.org/10.1161/CIR.0000000000000973.
60. Via M. The malnutrition of obesity: micronutrient deficiencies that promote diabetes. ISRN Endocrinology 2012;103472. https://doi.org/10.5402/2012/103472.
61. Brown T, Mann B, Ryder N, Subbiah M, Kaplan JD, Dhariwal P, et al. Language models are few-shot learners. Adv Neural Inf Process Syst. 2020;33:1877–901.
62. Raffel C, Shazeer N, Roberts A, Lee K, Narang S, Matena M, et al. Exploring the limits of transfer learning with a unified text-to-text transformer. J Mach Learn Res. 2020;21:1–67.
63. Roberts RH, Ali SR, Hutchings HA, Dobbs TD, Whitaker IS. Comparative study of ChatGPT and human evaluators on the assessment of medical literature according to recognised reporting standards. BMJ Health Care Inform 2023;30. https://doi.org/10.1136/bmjhci-2023-100830.
64. Milmo D. ChatGPT reaches 100 million users two months after launch. Guardian 2023;3. https://www.theguardian.com/technology/2023/feb/02/chatgpt-100-million-users-open-ai-fastest-growing-app.
65. OpenAI. ChatGPT. Optimizing language models for dialogue. https://openai.com/blog/chatgpt. Accessed 14 December 2023.
66. Reiss MV. Testing the reliability of chatgpt for text annotation and classification: A cautionary remark. arXiv preprint arXiv:230411085 2023. https://osf.io/preprints/osf/rvy5p.
67. Meskó B. Prompt engineering as an important emerging skill for medical professionals: tutorial. J Med Internet Res. 2023;25:e50638. https://doi.org/10.2196/50638.

AUTHOR CONTRIBUTIONS

FN contributed to the conceptualization, methodology, writing - original draft preparation, project supervision and administration. MT, DM, SK, AM and MA contributed to the formal analysis. BU contributed to the investigation and data curation. LN contributed to the conceptualization, writing - original draft, reviewing and editing, project supervision and administration. All authors reviewed and commented on subsequent drafts of the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41430-024-01476-y.

Correspondence and requests for materials should be addressed to Lara Nasreddine.