

Review

From AI to the Table: A Systematic Review of ChatGPT's Potential and Performance in Meal Planning and Dietary Recommendations

Peiqi Guo ^{1,*} , Guancheng Liu ², Xiaoling Xiang ³ and Ruopeng An ⁴¹ School of Public Health, University of Texas Health Science Center at Houston, Houston, TX 77030, USA² School of Sport and Exercise Science, University of Northern Colorado, Greeley, CO 80639, USA; guancheng.liu@unco.edu³ School of Social Work, University of Michigan, Ann Arbor, MI 48109, USA; xiangxi@umich.edu⁴ Silver School of Social Work, New York University, New York, NY 10003, USA; ruopeng.an@nyu.edu

* Correspondence: peiqi.guo@uth.tmc.edu

Abstract: A balanced diet is crucial for preventing diseases and managing existing health conditions. ChatGPT has garnered attention from researchers, including nutrition scientists and dietitians, as an innovative tool for personalized meal planning and dietary recommendations. Objectives: The purpose of this study was to review scientific evidence on ChatGPT's performance in providing personalized meal plans and generating dietary recommendations. Methods: This systematic review was conducted following the PRISMA guidelines. Keyword-based database searches were performed on PubMed, Web of Science, EBSCO, and Embase. Inclusion criteria included (1) empirical studies and (2) primary research on ChatGPT's performance in personalized meal planning and dietary recommendations. Results: Twenty-three studies met the inclusion criteria, comprising fourteen validation studies, five comparative studies, and four qualitative studies. Most studies reported that ChatGPT achieved satisfactory accuracy and was often indistinguishable from human dietitians. One study even reported that ChatGPT outperformed human dietitians. However, limitations and risks, such as safety concerns and a lack of real-world implementation, were also identified. Conclusions: ChatGPT shows promise as a relatively reliable innovative tool for personalized meal planning and dietary recommendations, offering more accessible and cost-effective solutions. Nevertheless, further studies are needed to address its limitations and challenges.

Keywords: ChatGPT; personalized dietary recommendations; AI-driven nutrition



Academic Editor: Dalia El Khoury

Received: 1 December 2024

Revised: 6 January 2025

Accepted: 5 February 2025

Published: 14 February 2025

Citation: Guo, P.; Liu, G.; Xiang, X.; An, R. From AI to the Table: A Systematic Review of ChatGPT's Potential and Performance in Meal Planning and Dietary Recommendations. *Dietetics* **2025**, *4*, 7. <https://doi.org/10.3390/dietetics4010007>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

A balanced diet is fundamental to human health, influencing the risks and development of numerous health conditions [1]. Studies have shown that poor dietary behaviors are associated with various health conditions, including but not limited to, hypertension, type 2 diabetes, obesity, cardiovascular diseases (CVD), and certain types of cancer. Excessive sodium intake, for instance, is a known contributor to hypertension [2,3]. Diets high in added sugars are linked to type 2 diabetes and obesity [4], while high consumption of unhealthy fats (such as saturated and trans fats), cholesterol, and sodium has been shown to increase the risk of CVD [5]. Additionally, consuming foods containing carcinogens or diets that are themselves carcinogenic, such as those involving alcohol, raises the risk of several types of cancers [6]. These chronic diseases affect millions of people in the

United States. For example, data from the Centers for Disease Control and Prevention (CDC) indicate that, in 2017–2018, approximately 42.4% of American adults were obese [7]. Fortunately, evidence-based research has consistently shown that a healthy diet brings significant benefits to chronic disease treatment and management [8,9]. In addition, even healthy populations can benefit from proper diet and nutrition management by preventing illnesses. Therefore, dietary and nutrition management are essential components of modern health management.

While the benefits of a balanced diet in disease treatment are widely recognized, personalized dietary management elevates these effects further. Personalized dietary recommendations, which are tailored to individual factors such as age, sex, activity level, and medical history, have been shown to result in better adherence and favorable health outcomes compared to generalized dietary guidelines. For example, studies have found that personalized diets can reduce HbA1c levels, a key indicator of blood glucose in type 2 diabetes patients, by 0.5–1% [10]. These findings highlight the potential of personalized dietary recommendations in clinical practice and real-world applications.

However, access to personalized dietary services is often restricted due to significant barriers. Insurance coverage for nutritional counseling is typically limited to specific medical conditions, such as diabetes or eating disorders, and is not widely available for preventive care. For instance, registered dietitians are not recognized as Medicaid providers in some states, and nutrition counseling is not a standardized covered service under Medicaid. Medicare provides coverage only for beneficiaries with specific conditions. This variability in coverage creates substantial cost barriers, making these services inaccessible to many individuals, particularly underserved populations, and contributing to health disparities. Additionally, personalized dietary recommendations require comprehensive individual assessments and multidisciplinary expertise, which pose practical challenges to implementation and scalability [11,12].

The prevalence of online searches for health-related information, including serious medical advice, has significantly increased in the Internet era [13]. At the same time, information technologies have rapidly evolved over the past few decades. Advances in Natural Language Processing (NLP) have accelerated the development of large language models (LLMs). Among these, Chat Generative Pre-Trained Transformer (ChatGPT), developed by OpenAI and introduced to the public in 2020, stands out as one of the most well-known and widely used LLMs. ChatGPT features advanced NLP capabilities, allowing it to generate highly natural, human-like responses surpassing many existing chatbots [14–16]. Its accessibility, tailored responses, human-like output, and exceptional performance have garnered attention from researchers across diverse fields, including nutrition science and dietetics.

Innovative tools like ChatGPT and other AI-driven LLMs offer a promising solution to access barriers to personalized nutritional counseling. These technologies provide a cost-effective, scalable, and versatile alternative to traditional consultations, enabling broader access to dietary guidance. By delivering personalized, evidence-based recommendations, tools like ChatGPT can complement traditional services and bridge gaps in nutritional support, particularly for populations with limited access to registered dietitians. With careful design and oversight, these tools have the potential to enhance health equity by making high-quality dietary services accessible to a much broader audience, addressing disparities, and promoting better overall health outcomes.

Researchers have begun exploring ChatGPT's potential, capabilities, and effectiveness in addressing dietary issues such as common nutrition inquiries, personalized diet recommendations, dietary plans, and food identification. However, individual studies often vary widely in their methodologies and frequently reach mixed or inconclusive conclusions [17]. This inconsistency makes assessing ChatGPT's overall performance challenging and limits

the development of clear, evidence-based guidance on its applications in nutrition. Furthermore, key challenges persist, including ensuring the safety of dietary recommendations and achieving consistent performance across diverse contexts and health conditions.

Given the growing interest in AI and LLM applications in dietary and nutrition sciences and the increasing number of studies in this area, this systematic review is timely and necessary. This review aims to address the existing gaps by systematically evaluating evidence on ChatGPT's performance in personalized meal planning and dietary recommendations. By critically analyzing findings from 23 studies [18–40], this study seeks to provide a comprehensive assessment of current evidence, including the quality of methodologies, strengths, and limitations. Such an evaluation is essential for developing a clearer understanding of the state of the evidence, guiding future research directions, and informing the integration of AI tools like ChatGPT into dietary services.

2. Methods

2.1. Study Selection Protocols

This review followed the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) guideline [41]. Studies were included if they met all the following criteria: (1) the paper is a primary empirical study, not a review article; (2) the paper focuses on the application of ChatGPT in personalized meal planning and dietary recommendation; (3) the paper is published in English. The exclusion criteria were as follows: (1) papers published before 2022, when ChatGPT was introduced to the public; (2) studies evaluating other LLMs without including ChatGPT; (3) secondary articles such as reviews, commentaries, or editorials.

2.2. Search Strategy

The literature search for this review was conducted in October 2024 across five databases: PubMed, Scopus, Web of Science, EBSCO, and Embase. Search prompts were designed using a combination of keywords and restrictions, including Title and Abstract (tiab) and Medical Subject Headings (MeSH). For PubMed, the following prompt were used: ("ChatGPT") AND (diet[tiab] OR dietary[tiab] OR meal[tiab] OR meals[tiab] OR nutrition[tiab] OR nutritional[tiab] OR food[tiab] OR eating[tiab] OR menu[tiab] OR feeding[tiab] OR meals[mh] OR diet[mh] OR "diet therapy"[mh] OR "menu planning"[mh] OR "nutrition therapy"[mh] OR "feeding behavior"[mh] OR dietetics[mh] OR food[mh] OR eating[mh]). This prompt included keywords relevant to the review topics. Searching prompts for the other databases were adapted from this prompt using the same set of keywords and can be found in the Appendix A.

2.3. Data Extraction

Data extraction was performed by one author (PG) and subsequently verified for accuracy by two additional authors (GL and RA). A standardized data extraction form was used to extract and organize information from the included articles. The key features extracted were the first author's last name, year of publication, region of the study, diet suggestion category(s), ChatGPT model version, model customization status, health condition emphasis, key results, and study limitations. Meta-analyses were not conducted due to the heterogeneity in study results, health condition focus, and methodologies.

2.4. Study Quality Assessment

The study quality assessment criteria were adapted from previous research [42,43]. The quality of each included study was evaluated based on the following criteria: (1) clear research objectives and goals directly related to ChatGPT; (2) detailed description of the

sample; (3) specification of the ChatGPT model version used; (4) use of valid outcome measures; (5) comparison with an appropriate benchmark; (6) involvement of experts as references to evaluate AI outputs; (7) detailed description of study methods to ensure transparency in data collection; (8) transparent discussion of study limitations; and (9) consideration of real-world applicability. The quality of each study was assessed by presence (score = 1) or absence (score = 0). The final quality score, ranging from 0 to 8, was calculated by summing the points across all criteria.

3. Results

3.1. Study Selection

A total of 507 articles were identified through the database keyword search. Of these, 355 were excluded during the pre-screening phase for reasons such as duplication, being non-English articles, not being peer-reviewed, or not being empirical research articles. The remaining 152 articles were screened by title and abstract, resulting in the exclusion of 111 articles based on the exclusion criteria. The remaining 41 articles were further evaluated in full text against the study selection criteria. Among these, four were excluded for not being primary empirical studies; seven did not focus on personalized diet planning and recommendations; two evaluated other LLMs without including ChatGPT, and four utilized LLMs but did not address personalized dietary recommendations. Ultimately, 23 studies met the inclusion criteria and were included in the final review (Figure 1).

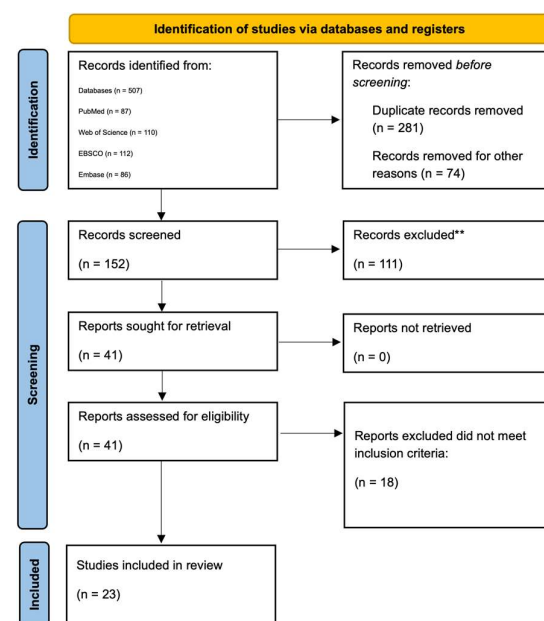


Figure 1. PRISMA Diagram.

3.2. Basic Characteristics of Included Studies

Table 1 summarizes the 23 studies included in the review. These comprise fifteen validation studies (Wang LC et al. [40]; Ponzo V et al. [31]; Tsai Ch et al. [38]; Sun H et al. [37]; Papastratis I et al. [28,29]; Niszczota P et al. [25]; Naqvi HA et al. [24]; Naja F et al. [23]; Liao LL et al. [19]; Kiriakedis S et al. [16]; Haman M et al. [13]; Aiumtrakul N et al. [3]; Acharya PC et al. [1]; Dimitriadis F et al. [7]), five comparative studies (Qarajeh A et al. [32]; Kirk D et al. [17]; Hieronimus B et al. [14]; Bayram HM et al. [5]; Agne I et al. [2]), and four qualitative studies providing descriptive insights (Lo FPW et al. [20]; Kim DW et al. [15]; Leslie-Miller CJ et al. [18]). We classified studies as “validation” studies if they specifically focused on directly assessing ChatGPT’s performance in providing personalized dietary

recommendations. In contrast, comparative studies evaluated ChatGPT's performance by comparing it to other LLMs or human dietitians.

Table 1. Basic Characteristics.

Study ID	First Author, Year	Region	Diet Suggestion Category	ChatGPT Model/Version Used	Customize/ Fine-tuned?	Health Condition Focused?
18	Wang, 2024	USA	Dietary Management, Nutritional Recommendation	GPT-4	No	Dialysis
19	Ponzo, 2024	Italy	Dietary Management	GPT-3.5	No	1. Dyslipidemia (hypercholesterolemia and hypertriglyceridemia) 2. Arterial hypertension 3. Type 2 diabetes mellitus (T2DM) 4. Obesity 5. Non-alcoholic fatty liver disease (NAFLD) 6. Chronic kidney disease (CKD) 7. Sarcopenia
20	Tsai, 2023	USA	Dietary Management	GPT-3.5	Customized APP	Pregnancy
21	Sun, 2023	China	Dietary Management	ChatGPT, GPT 4.0	No	T2DM
22	Qarajeh, 2023	USA, Jordan, Thailand	Nutrition Estimation	GPT-4, GPT-3.5	No	Chronic Kidney Disease (CKD)
23	Papastratis, 2024	Greece	Dietary Management	GPT-4, GPT-3.5	No	1. Obesity 2. Cardiovascular disease
24	Papastratis, 2024	Greece	Dietary Management	GPT-4	No	3. T2DM
25	Niszczoła, 2023	Poland	Dietary Management	GPT-3	No	No
26	Naqvi, 2024	USA	Dietary Management	GPT-3.5	No	Allergies Inflammatory Bowel Disease (IBD)
27	Naja, 2024	UAB, Lebanon, Bahrain	Dietary Management, Nutrition Estimation	GPT-3	No	1. T2DM 2. MetS 3. Hyperglycemia 4. Obesity 5. HTN 6. High TG 7. Low HDL Levels
28	Liao, 2024	Taiwan	Dietary Management	GPT-3.5	No	No
29	Leslie-Miller, 2024	USA	Dietary Management	N/A	No	Pediatric
30	Kirk, 2023	Netherlands	Dietary Management	GPT-3	No	No
31	Kiriakidis, 2024	USA	Dietary Management	GPT-4	No	Nephrolithiasis
32	Kim, 2024	USA, South Korea	Dietary Management	GPT-4	No	Multiple Health Conditions
33	Lo, 2024	UK, Hong Kong	Dietary Management, Nutrition Estimation	GPT-4v	No	NO
34	Hieronimus, 2024	Germany	Nutrition Recommendation	GPT-4	No	No
35	Haman, 2023	Czech	Nutritional Recommendation	GPT-3.5	No	No
36	Bayram, 2024	Turkey	Nutritional Recommendation	GPT-4, GPT-3.5	No	No
37	Aiumtrakul, 2024	USA, Thailand	Diet Management	GPT-4, GPT-3.5	No	Kidney Stone
38	Agne, 2024	Germany	Dietary Management, Nutritional Recommendation	GPT-3.5	No	Obesity
39	Acharya, 2024	USA, Thailand, Hungary	Dietary Management, Nutritional Recommendation	GPT-4, GPT-3.5	No	Chronic Kidney Disease (CKD)
40	Dimitriadis, 2024	Greece, Poland	Dietary Management	N/A	No	Heart Failure (HF)

In terms of particular health condition focus, 16 studies examined ChatGPT's performance in addressing one or more specific health conditions (Wang et al. [40]; Ponzo et al. [31]; Tsai Ch et al. [38]; Sun H et al. [37]; Qarajeh A et al. [32]; Papastratis I et al. [28]; Niszczoła P et al. [25]; Naqvi HA et al. [24]; Naja F et al. [23]; Leslie-Miller et al. [18]; Kiriakidis S et al. [16]; Kim DW et al. [15]; Aiumtrakul N et al. [3]; Agne I et al. [2]; Acharya PC et al. [1]; Dimitriadis F et al. [7]). The remaining nine studies explored ChatGPT's applications in general dietary advice and management.

3.3. ChatGPT Effectiveness on Diet Plan

Table 2 reports the results and conclusions from 23 studies that evaluated ChatGPT's performance in personalized diet recommendations across various contexts and health conditions. Most studies demonstrated that ChatGPT-generated dietary advice was highly rated and aligned with nutritional guidelines. However, areas such as calorie estimation, macronutrient balance, and adherence to clinical requirements require further improvement. Key strengths of ChatGPT's responses included: (a) accurate identification of food items, including cross-cultural foods (87.5% food detection accuracy); (b) alignment with credentialed nutrient information databases; and (c) generation of diverse meal plans [34].

Several studies highlighted that ChatGPT's meal plans often met recommended macronutrient and micronutrient requirements [23,24,32,34], with an average macronutrient accuracy of 84.19%. However, performance declined in complex scenarios involving coexisting health conditions and dietary restrictions for allergies. For example, ChatGPT-generated meal plans included allergens like nuts for individuals with allergies [29]. Challenges also included inconsistencies in accuracy for specific nutrient categories (e.g., high oxalate content) and variability in portion sizes and micronutrient estimates [31].

Comparative studies showed that GPT-4 outperformed GPT-3.5 in nutrient accuracy and meal variety [36]. While ChatGPT models often received positive ratings from professionals and, in some cases, outperformed human dietitians [25,28,30], incomplete or discordant dietary suggestions were occasionally identified. Experts emphasized the need for supervision of AI-generated content to avoid misunderstandings or misuse [25,35]. Figure 2 showed a comparison chart between GPT-4 and GPT-3.5.

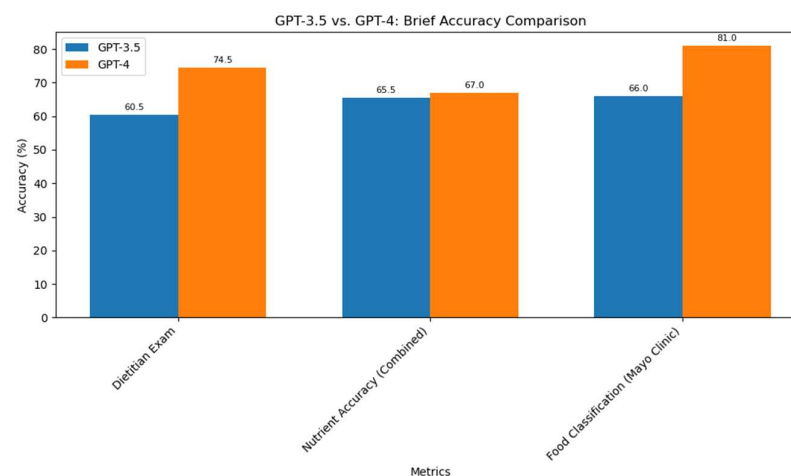


Figure 2. GPT-3.5 versus GPT-4 Brief Accuracy Comparison.

Real-world applications revealed mixed results. While ChatGPT performed well in providing general dietary recommendations aligned with clinical guidelines, its accuracy decreased when addressing complex dietary restrictions involving multiple health conditions and micronutrient needs. Studies emphasized the importance of further model training and refinement based on real-world guidelines and practical experiences. Furthermore, one study reported that 53% of experts believed ChatGPT's advice should not be solely relied upon for critical dietary plans or decisions [19].

Table 2. Study Results and Limitations.

Category	Study ID	First Author, Year	Key Results	Limitations
Validation Study	18	Wang, 2024	Renal dietitian rated ChatGPT generated meal plan as 5, and nutritional as 2 per 5-point Likert scale (low 1, high 5).	N/A
	19	Ponzo, 2024	Overall accuracy of ChatGPT's advice ranged from 55.5% (sarcopenia) to 73.3% (NAFLD).	Not performed on the most recent versions of ChatGPT.
	20	Tsai, 2023	A ChatGPT-powered chatbot was introduced.	N/A
	21	Sun, 2023	1. ChatGPT: 60.5% accuracy on dietitian exam. 2. GPT-4.0: 74.5% accuracy on the dietitian exam. 3. Ketogenic diet adherence: 80.7% (non-recommended foods), 94.87% (recommended foods).	1. Incomplete exposure to patient questions. 2. Variability in response. 3. Scope definition unclear.
	24	Papastratis, 2024	1. Energy intake deviation: 17%. 2. Macronutrient accuracy improves by 12%. 3. Overall accuracy: 84.19%.	N/A
	25	Niszczoła, 2023	1. 52/56 ChatGPT-generated diets include allergens. 2. Frequent errors in food quantity specification.	More dynamic interactions with ChatGPT.
	26	Naqvi, 2024	Appropriate response rate: 83.3%. Inter-rater reliability: 94.4%.	GPT's limitations in data interpretation.
	27	Naja, 2024	1. Incomplete/discordant dietary management recommendations. 2. Lacks complete PES statements in nutrition care. 3. Diet plan: display micronutrients and macronutrients discrepancies.	1. Same prompt vary response. 2. Differences bring by prompt quality. 3. Lack of human comparison group.
	28	Liao, 2024	1. 84.38% accuracy rate Nutrition Literacy (NL) test. 2. Feedback: 'Lacks thoroughness/rigor,' cited 52 times among 30 dietitians' 242 entries.	1. Restrictive scenarios may not represent all dietary challenges. 2. Potential bias introduced by misunderstood dietitians. 3. Small sample size.
	31	Kiriakedis, 2024	ChatGPT aligns dietary recommendations with clinical guidelines.	ChatGPT's recommendations may not account for the personalized situation of patients.
	32	Kim, 2024	1. 5/14 experts successfully distinguish AI from human content. 2. 79.1% (53/67) of experts unable to distinguish AI-generated diet plans, rated similarly to controls. 3. The AI-generated diet plan was rated above neutral in all evaluation variables.	The AI diet plan has several limitations: 1. Conflicts about dietary considerations. 2. Insufficient details about recommendations. 3. Lack of affordability.
	33	Lo, 2024	1. 87.5% accuracy in food detection. 2. Limited performance in portion size estimations. 3. Nutritional contents conversion is well-aligned with the USDA National Nutrient Database. 4. Able to identify regional dishes.	1. Portion Size estimation for small size. 2. Limited performance in portion estimation.
	34	Hieronimus, 2024	1. ChatGPT-generated meal plan mostly met the nutrient requirement of macronutrients and micronutrients.	1. Individual user prompts may influence the result. 2. Quick evolving on GPT models suggests the result is only valid for study time.
	35	Haman, 2023	97% of values fall within the 40% difference in USDA data.	1. Restricted to basic nutrition of a single item. 2. Accuracy variable across different nutrients. 3. Unable to assess ChatGPT Plan for chronic conditions patients.
	36	Bayram, 2024	GPT 4 achieved 60.6% accuracy for the food classification task.	1. Limited in experimental environment. 2. Accuracy discrepancies exist between different models.
Comparative Study	37	Aiumtrakul, 2024	1. Accuracy: GPT-4 (52%), GPT-3.5(49%). 2. Accuracy decreases with higher oxalate content categories.	3. Only focused on PQ classification. 1. Enhancement in the chatbot algorithm is needed for better clinical applicability. 2. The study only focuses on oxalate content.
	39	Acharya, 2024	1. Both GPT-3.5 and GPT-4 provide no misleading response to all questions of KDOQI and KDIGO guideline questions. 2. Flesch-Kincaid Grade Level readability assessment: ChatGPT 3.5 (11.3 ± 2.1), ChatGPT 4.0 (11.1 ± 1.9).	1. Evaluation criteria materials may not represent the entirety and complexity of CKD patients. 2. Absent of need for further assessment based on real-world inquiries.
	40	Dimitriadis, 2024	ChatGPT provided thorough and accurate answers for 7/8 questions.	1. The study does not involve real patients. 2. Only assess HF and accuracy performance.
	22	Qarajeh, 2023	1. GPT-4: 81% accuracy, GPT-3.5: 66% (Mayo Clinic Renal Diet, 240 items). 2. Potassium: GPT-4 (81%), GPT-3.5 (66%). 3. Phosphorus: GPT-4 (77%), GPT-3.5 (85%).	1. Lack of real-world food diversity. 2. Model discrepancies remain unexplored. 3. Lack of consideration for CKD. 4. Leaving potential concerns such as user experience, usability, and AI integration with clinical workflow.
	23	Papastratis, 2024	1. Nutrient accuracy: GPT-3.5 (82%), GPT-4 (82%), KB Recommender (91%). 2. Energy intake accuracy (caloric difference): ChatGPT (>19%), KB Recommender (0.8%). 3. Meal variety score: GPT-3.5 (6.58), GPT-4 (6.4), KB Recommender (4.89).	N/A
	29	Leslie-Miller, 2024	No difference was observed between ChatGPT and Expert.	1. Self-report bias 2. Restricted populations 3. Lack of long-term impact measuring
	30	Kirk, 2023	ChatGPT receives higher overall grades than dietitians on five occasions.	The limit of response length may limit response quality.
	38	Agne, 2024	1. ChatGPT provides suitable personalized dietary advice compared with another algorithm. 2. ChatGPT recommendations contain inconsistencies. 3. ChatGPT should not be solely relied upon.	1. Inconsistencies and errors in diet management and recommendations 2. ChatGPT response lack of depth and reliability. 3. Further development and fine-tuning of LLM models are necessary to fit dietary recommendations needs.

3.4. Study Quality

Table 3 reports the results of the study's quality assessment. On average, the included studies met seven out of nine quality criteria. All but two studies clearly stated research objectives and goals directly related to ChatGPT applications in personalized dietary recommendations [29,40]. Nineteen articles clearly described the samples or participants involved in the study [18–31,33–37]. Twenty studies specified the ChatGPT model version used, while three did not [21,29,40]. Benchmark comparison received the lowest score, with only 15 studies comparing ChatGPT's performance against appropriate benchmarks [18–20,22–26,28,30,31,33,36–38]. Additionally, 18 studies involved experts as references to evaluate AI outputs, while five did not [21,27,29,32,39]. Finally, 17 studies discussed their limitations and real-world applicability, while six did not [21,27,29,32,35,39].

Table 3. Quality Assessment of Studies.

Item	Criterion of Study Quality	Percentage Met Criteria/Mean
1	Clear research objectives and goals directly related to ChatGPT.	91%
2	Detailed description of the sample.	83%
3	Specification of the ChatGPT model version used.	87%
4	Valid outcome measures.	91%
5	Comparison with an appropriate benchmark.	65%
6	Involvement of experts as references to evaluate AI outputs.	78%
7	Detailed description of study methods to ensure transparency in data collection.	91%
8	Transparent discussion of study limitations and consideration of real-world applicability.	78%
9	Total study quality score by summing up items 1–8.	7.3

4. Discussion

Dietary and nutritional management are increasingly important in health promotion and disease prevention. Advances in AI have introduced innovative possibilities in the dietary field, with tools like ChatGPT showing significant potential. This systematic review evaluated existing scientific evidence on ChatGPT's performance in personalized diet recommendations and nutritional management. A total of 23 studies met the inclusion criteria identified through comprehensive database searches and rigorous screening processes. Among these, 20 studies reported positive results, highlighting ChatGPT's ability to perform various tasks related to dietary planning. ChatGPT demonstrated strengths in providing dietary advice aligned with clinical guidelines, excelling in areas such as food identification, nutrient accuracy, and meal variety. Studies also indicated that many users and experts could not distinguish ChatGPT-generated diet plans from those created by human dietitians. These findings suggest that ChatGPT holds great promise for enhancing dietary services.

Comparative studies included in this review explored ChatGPT's performance across three main dimensions: (1) version comparisons involving comparing different versions of ChatGPT (e.g., GPT-3.5 vs. GPT-4). These studies revealed that newer versions generally performed better in nutrient accuracy and meal variety. (2) ChatGPT versus human dietitians: studies comparing ChatGPT to human dietitians found that ChatGPT often produced recommendations comparable to those of professionals, though expert feedback highlighted the need for human supervision to address incomplete or inconsistent outputs. (3) ChatGPT versus benchmarks: comparisons against certified nutritional databases or official guidelines provided quantifiable insights into ChatGPT's intrinsic performance, although inconsistent benchmarks across studies made it challenging to unify findings.

4.1. Limitations

This study focuses on ChatGPT as the Large Language Model (LLM) and does not include other LLMs for several reasons. First, ChatGPT is one of the most dominant LLMs in the market and is widely studied in academia, serving as an exemplary LLM. In addition, there are numerous LLMs available, making it impossible to assess all their performances. Thus, the study focuses only on ChatGPT. However, it would be beneficial to further study the performance differences across various LLMs in dietary recommendations and related applications.

This review identified several limitations in the extant evidence. Over half of the included studies were conducted in experimental or simulated settings, lacking observational data from real-world use cases. This limits the generalizability of findings. Fourteen studies focused on specific health conditions, restricting their findings to narrow fields. Broader investigations encompassing general dietary applications are needed to further assess ChatGPT's performance across diverse real-world scenarios. None of the included studies examined ChatGPT's long-term impact on dietary adherence and health outcomes. Intervention-based research designs involving human subjects in real-world settings are needed to evaluate the sustained effects of ChatGPT-generated dietary plans.

Significant heterogeneity existed in the methodologies employed across the included studies. A key challenge in assessing ChatGPT's performance is determining the appropriate benchmark. Some studies used clinical guidelines from authoritative institutions like the NIH, while some relied on ratings from experts such as Registered Dietitians [18–20,33,36]. In addition, inconsistent prompt designs were used [23,24,32], which likely introduced biases in performance outcomes, further complicating the comparability of findings across studies. This heterogeneity in study designs and methodologies precluded the possibility of performing a meta-analysis for this review.

Challenges and inconsistencies in ChatGPT's performance were also identified. A major concern is nutritional safety. For instance, one study focusing on allergy-related dietary recommendations found that ChatGPT's meal plans included allergens, posing risks for individuals with specific food allergies [25]. Considering that these populations are less tolerant to nutritional imbalances compared to healthy individuals, inconsistent nutritional recommendations may adversely affect their health conditions and worsen their situations. Additionally, ChatGPT's performance was inconsistent and less reliable in managing complex cases involving coexisting health conditions or highly restrictive dietary needs. While ChatGPT may perform well on dietitian exams and receive positive ratings from dietitians, it delivers unsatisfactory performance on restricted meal suggestions, such as those involving allergens [21,25,28,30].

It is important to acknowledge an inherent bias in the use of ChatGPT, as its output is highly dependent on the quality and format of the input provided. Variations in input types and user expertise in using ChatGPT can lead to differences in the quality of dietary recommendation results. Users with limited knowledge of ChatGPT or inadequate skills in prompt design may be unable to fully utilize its advanced functionalities. Further research is beneficial to investigate how user familiarity with ChatGPT influences the quality and performance of its nutritional recommendations.

4.2. Broader Implications

Despite the limitations, the review highlights ChatGPT's potential as a tool for nutrition management and personalized diet recommendations. ChatGPT offers several advantages, including accessibility, cost-effectiveness, and versatility. Its accessibility stems from its digital format, which allows users to access it through online platforms or mobile applications at any time. This 24/7 availability ensures that users can interact with ChatGPT

as long as they have an internet-connected device. Compared to the cost of consultations with human dietitians, ChatGPT represents a more cost-effective alternative for providing dietary advice [28]. Additionally, ChatGPT has demonstrated versatility in addressing diverse dietary needs. While its performance diminishes when managing cases involving multiple disease conditions, it has shown competence in performing various roles traditionally associated with specialized dietitians, such as providing general nutrition advice or clinical dietary guidance for specific illnesses. These attributes suggest that ChatGPT could complement human dietitians and integrate into existing healthcare systems as an auxiliary tool, provided its outputs are carefully supervised by experts.

Several promising directions for future research emerged from this review. First, developing standardized methods for evaluating the performance of ChatGPT and other LLMs in nutritional science will enhance the reliability and comparability of findings across studies. Second, further research is required to observe ChatGPT's performance in real-world dietary planning and nutrition management scenarios, moving beyond the limitations of controlled experimental settings. Third, future studies should investigate the impact of prompt engineering on ChatGPT's outputs in the context of dietary recommendations. Standardizing prompt design could help reduce response variability and improve the consistency of outputs. Fourth, investigating the long-term effects of ChatGPT-generated dietary plans on adherence and health outcomes is crucial. Finally, studies should examine the feasibility of integrating ChatGPT into healthcare services as a collaborative tool for dietitians and physicians. Addressing these research gaps will advance our understanding of ChatGPT's potential in nutrition management and dietary planning, ultimately making dietary services more accessible and widely used by patients.

5. Conclusions

Personalized meal planning and dietary recommendations are essential for health promotion and disease prevention across diverse populations. However, access to dietary services remains limited due to cost and other access barriers. This systematic review highlights ChatGPT's potential as an innovative tool for personalized nutrition. The findings demonstrate that ChatGPT can provide accurate and effective dietary advice, though challenges related to consistency, safety, and real-world applications persist. Addressing these limitations through future research will enable ChatGPT to serve as a complementary tool within healthcare systems, expanding access to reliable nutrition solutions and improving overall health and well-being for diverse and often marginalized communities.

Author Contributions: Conceptualization, R.A.; formal analysis, P.G. and G.L.; writing—original draft preparation, P.G.; writing—review and editing, P.G., G.L., X.X. and R.A.; supervision, R.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Search Prompt

1. PubMed

Prompt:

("ChatGPT") AND (diet[tiab] OR dietary[tiab] OR meal[tiab] OR meals[tiab] OR nutrition[tiab] OR nutritional[tiab] OR food[tiab] OR eating[tiab] OR menu[tiab] OR feeding[tiab] OR meals[mh] OR diet[mh] OR "diet therapy"[mh] OR "menu planning"[mh] OR "nutrition therapy"[mh] OR "feeding behavior"[mh] OR dietetics[mh] OR food[mh] OR eating[mh])

2. Web of Science

Prompt:

“ChatGPT” AND (diet OR dietary OR meal OR meals OR nutrition OR nutritional OR food OR eating OR menu OR feeding OR diet OR “diet therapy” OR “menu planning” OR “nutrition therapy” OR “feeding behavior” OR dietetics OR food OR eating) (Title) or “ChatGPT” AND (diet OR dietary OR meal OR meals OR nutrition OR nutritional OR food OR eating OR menu OR feeding OR diet OR “diet therapy” OR “menu planning” OR “nutrition therapy” OR “feeding behavior” OR dietetics OR food OR eating) (Abstract)

3. EBSCO

Prompt:

TI (“ChatGPT” AND (diet OR dietary OR meal OR meals OR nutrition OR nutritional OR food OR eating OR menu OR feeding OR diet OR “diet therapy” OR “menu planning” OR “nutrition therapy” OR “feeding behavior” OR dietetics OR food OR eating)) OR AB (“ChatGPT” AND (diet OR dietary OR meal OR meals OR nutrition OR nutritional OR food OR eating OR menu OR feeding OR diet OR “diet therapy” OR “menu planning” OR “nutrition therapy” OR “feeding behavior” OR dietetics OR food OR eating))

4. Embase

Prompt:

“ChatGPT” AND (diet OR dietary OR meal OR meals OR nutrition OR nutritional OR food OR eating OR menu OR feeding OR diet OR “diet therapy” OR “menu planning” OR “nutrition therapy” OR “feeding behavior” OR dietetics OR food OR eating)

References

1. Acharya, P.C.; Alba, R.; Krisanapan, P.; Acharya, C.M.; Suppadungsuk, S.; Csongradi, E.; Mao, M.A.; Craici, I.M.; Miao, J.; Thongprayoon, C.; et al. AI-Driven Patient Education in Chronic Kidney Disease: Evaluating Chatbot Responses against Clinical Guidelines. *Diseases* **2024**, *12*, 185. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Agne, I.; Gedrich, K. Personalized dietary recommendations for obese individuals—A comparison of ChatGPT and the Food4Me algorithm. *Clin. Nutr. Open Sci.* **2024**, *56*, 192–201. [\[CrossRef\]](#)
3. Aiumtrakul, N.; Thongprayoon, C.; Arayangkool, C.; Vo, K.B.; Wannaphut, C.; Suppadungsuk, S.; Krisanapan, P.; Garcia Valencia, O.A.; Qureshi, F.; Miao, J.; et al. Personalized Medicine in Urolithiasis: AI Chatbot-Assisted Dietary Management of Oxalate for Kidney Stone Prevention. *J. Pers. Med.* **2024**, *14*, 107. [\[CrossRef\]](#)
4. Alawida, M.; Mejri, S.; Mehmood, A.; Chikhaoui, B.; Isaac Abiodun, O. A Comprehensive Study of ChatGPT: Advancements, Limitations, and Ethical Considerations in Natural Language Processing and Cybersecurity. *Information* **2023**, *14*, 462. [\[CrossRef\]](#)
5. Bayram, H.M.; Ozturkcan, A. AI showdown: Info accuracy on protein quality content in foods from ChatGPT 3.5, ChatGPT 4, bard AI and bing chat. *Br. Food J.* **2024**, *126*, 3335–3346. [\[CrossRef\]](#)
6. 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*, 1530–1531. [\[CrossRef\]](#)
7. Dimitriadis, F.; Alkagiet, S.; Tsigkriki, L.; Kleitsioti, P.; Sidiropoulos, G.; Efstratiou, D.; Askalidi, T.; Tsaousidis, A.; Siarkos, M.; Giannakopoulou, P.; et al. ChatGPT and Patients with Heart Failure. *Angiology* **2024**, 00033197241238403. [\[CrossRef\]](#)
8. Evert, A.B.; Dennison, M.; Gardner, C.D.; Garvey, W.T.; Lau, K.H.K.; MacLeod, J.; Mitri, J.; Pereira, R.F.; Rawlings, K.; Robinson, S.; et al. Nutrition Therapy for Adults with Diabetes or Prediabetes: A Consensus Report. *Diabetes Care* **2019**, *42*, 731–754. [\[CrossRef\]](#)
9. Forouhi, N.G.; Krauss, R.M.; Taubes, G.; Willett, W. Dietary fat and cardiometabolic health: Evidence, controversies, and consensus for guidance. *BMJ* **2018**, *361*, k2139. [\[CrossRef\]](#)
10. Garcia, M.B. ChatGPT as a Virtual Dietitian: Exploring Its Potential as a Tool for Improving Nutrition Knowledge. *Appl. Syst. Innov.* **2023**, *6*, 96. [\[CrossRef\]](#)
11. Go, V.L.W.; Wong, D.A.; Butrum, R. Diet, Nutrition and Cancer Prevention: Where Are We Going from Here? *J. Nutr.* **2001**, *131*, 3121S–3126S. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Hales, C.M.; Carroll, M.D.; Fryar, C.D.; Ogden, C.L. Prevalence of Obesity and Severe Obesity Among Adults: United States, 2017–2018. In *NCHS Data Brief*; National Center for Health Statistics: Hyattsville, MD, USA, 2020.
13. Haman, M.; Školník, M.; Lošťák, M. AI dietitian: Unveiling the accuracy of ChatGPT’s nutritional estimations. *Nutrition* **2024**, *119*, 112325. [\[CrossRef\]](#)

14. Hieronimus, B.; Hammann, S.; Podszun, M.C. Can the AI tools ChatGPT and Bard generate energy, macro- and micro-nutrient sufficient meal plans for different dietary patterns? *Nutr. Res.* **2024**, *128*, 105–114. [[CrossRef](#)]
15. Kim, D.W.; Park, J.S.; Sharma, K.; Velazquez, A.; Li, L.; Ostrominski, J.W.; Tran, T.; Seitter Pérez, R.H.; Shin, J.-H. Qualitative evaluation of artificial intelligence-generated weight management diet plans. *Front. Nutr.* **2024**, *11*, 1374834. [[CrossRef](#)]
16. Kiriakedis, S.; Duty, B.; Chase, T.; Wusirika, R.; Metzler, I. Using ChatGPT-4 to Analyze 24-Hour Urine Results and Generate Custom Dietary Recommendations for Nephrolithiasis. *J. Endourol.* **2024**, *38*, 719–724. [[CrossRef](#)]
17. Kirk, D.; van Eijnatten, E.; Camps, G. Comparison of Answers between ChatGPT and Human Dieticians to Common Nutrition Questions. *J. Nutr. Metab.* **2023**, *2023*, 5548684. [[CrossRef](#)]
18. Leslie-Miller, C.J.; Simon, S.L.; Dean, K.; Mokhallati, N.; Cushing, C.C. The critical need for expert oversight of ChatGPT: Prompt engineering for safeguarding child healthcare information. *J. Pediatr. Psychol.* **2024**, *49*, 812–817. [[CrossRef](#)]
19. Liao, L.-L.; Chang, L.-C.; Lai, I.-J. Assessing the Quality of ChatGPT's Dietary Advice for College Students from Dietitians' Perspectives. *Nutrients* **2024**, *16*, 1939. [[CrossRef](#)]
20. Lo, F.P.-W.; Qiu, J.; Wang, Z.; Chen, J.; Xiao, B.; Yuan, W.; Giannarou, S.; Frost, G.; Lo, B. Dietary Assessment with Multimodal ChatGPT: A Systematic Analysis. *IEEE J. Biomed. Health Inform.* **2024**, *28*, 7577–7587. [[CrossRef](#)]
21. Malik, V.S.; Hu, F.B. Sweeteners and Risk of Obesity and Type 2 Diabetes: The Role of Sugar-Sweetened Beverages. *Curr. Diabetes Rep.* **2012**, *12*, 195–203. [[CrossRef](#)]
22. Micha, R.; Peñalvo, J.L.; Cudhea, F.; Imamura, F.; Rehm, C.D.; Mozaffarian, D. Association Between Dietary Factors and Mortality From Heart Disease, Stroke, and Type 2 Diabetes in the United States. *JAMA* **2017**, *317*, 912–924. [[CrossRef](#)] [[PubMed](#)]
23. Naja, F.; Taktouk, M.; Matbouli, D.; Khaleel, S.; Maher, A.; Uzun, B.; Alameddine, M.; Nasreddine, L. Artificial intelligence chatbots for the nutrition management of diabetes and the metabolic syndrome. *Eur. J. Clin. Nutr.* **2024**, *78*, 887–896. [[CrossRef](#)] [[PubMed](#)]
24. Naqvi, H.A.; Delungahawatta, T.; Atarere, J.O.; Bandaru, S.K.; Barrow, J.B.; Mattar, M.C. Evaluation of online chat-based artificial intelligence responses about inflammatory bowel disease and diet. *Eur. J. Gastroenterol. Hepatol.* **2024**, *36*, 1109. [[CrossRef](#)] [[PubMed](#)]
25. Niszczota, P.; Rybicka, I. The credibility of dietary advice formulated by ChatGPT: Robo-diets for people with food allergies. *Nutrition* **2023**, *112*, 112076. [[CrossRef](#)] [[PubMed](#)]
26. Ohlhorst, S.D.; Russell, R.; Bier, D.; Klurfeld, D.M.; Li, Z.; Mein, J.R.; Milner, J.; Ross, A.C.; Stover, P.; Konopka, E. Nutrition research to affect food and a healthy lifespan. *Adv. Nutr.* **2013**, *4*, 579–584. [[CrossRef](#)]
27. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, n71. [[CrossRef](#)]
28. Papastratis, I.; Konstantinidis, D.; Daras, P.; Dimitropoulos, K. AI nutrition recommendation using a deep generative model and ChatGPT. *Sci. Rep.* **2024**, *14*, 14620. [[CrossRef](#)]
29. Papastratis, I.; Stergioulas, A.; Konstantinidis, D.; Daras, P.; Dimitropoulos, K. Can ChatGPT Provid. Approp. Meal Plans NCD Patients? *Nutrition* **2024**, *121*, 112291. [[CrossRef](#)]
30. Picazo-Sanchez, P.; Ortiz-Martin, L. Analysing the impact of ChatGPT in research. *Appl. Intell.* **2024**, *54*, 4172–4188. [[CrossRef](#)]
31. Ponzo, V.; Goitre, I.; Favaro, E.; Merlo, F.D.; Mancino, M.V.; Riso, S.; Bo, S. Is ChatGPT an Effective Tool for Providing Dietary Advice? *Nutrients* **2024**, *16*, 469. [[CrossRef](#)]
32. Qarajeh, A.; Tangpanithandee, S.; Thongprayoon, C.; Suppadungsuk, S.; Krisanapan, P.; Aiumtrakul, N.; Garcia Valencia, O.A.; Miao, J.; Qureshi, F.; Cheungpasitporn, W. AI-Powered Renal Diet Support: Performance of ChatGPT, Bard AI, and Bing Chat. *Clin. Pract.* **2023**, *13*, 104. [[CrossRef](#)] [[PubMed](#)]
33. Rahman, I.; Athar, M.T.; Islam, M. Type 2 Diabetes, Obesity, and Cancer Share Some Common and Critical Pathways. *Front. Oncol.* **2021**, *10*, 600824. [[CrossRef](#)] [[PubMed](#)]
34. Rosato, V.; Temple, N.J.; La Vecchia, C.; Castellan, G.; Tavani, A.; Guercio, V. Mediterranean diet and cardiovascular disease: A systematic review and meta-analysis of observational studies. *Eur. J. Nutr.* **2019**, *58*, 173–191. [[CrossRef](#)] [[PubMed](#)]
35. Singar, S.; Nagpal, R.; Arjmandi, B.H.; Akhavan, N.S. Personalized Nutrition: Tailoring Dietary Recommendations through Genetic Insights. *Nutrients* **2024**, *16*, 2673. [[CrossRef](#)]
36. Soltani, S.; Arablou, T.; Jayedi, A.; Salehi-Abargouei, A. Adherence to the dietary approaches to stop hypertension (DASH) diet in relation to all-cause and cause-specific mortality: A systematic review and dose-response meta-analysis of prospective cohort studies. *Nutr. J.* **2020**, *19*, 37. [[CrossRef](#)]
37. Sun, H.; Zhang, K.; Lan, W.; Gu, Q.; Jiang, G.; Yang, X.; Qin, W.; Han, D. An AI Dietitian for Type 2 Diabetes Mellitus Management Based on Large Language and Image Recognition Models: Preclinical Concept Validation Study. *J. Med. Internet Res.* **2023**, *25*, e51300. [[CrossRef](#)]

38. Tsai, C.-H.; Kadire, S.; Sreeramdas, T.; VanOrmer, M.; Thoene, M.; Hanson, C.; Berry, A.; Khazanchi, D. Generating Personalized Pregnancy Nutrition Recommendations with GPT-Powered AI Chatbot. In Proceedings of the 20th International Conference on Information Systems for Crisis Response and Management (ISCRAM), Omaha, NE, USA, 28–31 May 2023; pp. 263–271.
39. Verma, M.; Hontecillas, R.; Tubau-Juni, N.; Abedi, V.; Bassaganya-Riera, J. Challenges in Personalized Nutrition and Health. *Front. Nutr.* **2018**, *5*, 117. [[CrossRef](#)]
40. Wang, L.-C.; Zhang, H.; Ginsberg, N.; Nandorine Ban, A.; Kooman, J.P.; Kotanko, P. Application of ChatGPT to Support Nutritional Recommendations for Dialysis Patients—A Qualitative and Quantitative Evaluation. *J. Ren. Nutr.* **2024**, *34*, 477–481. [[CrossRef](#)]
41. Wong, D.K.-K.; Cheung, M.-K. Online Health Information Seeking and eHealth Literacy Among Patients Attending a Primary Care Clinic in Hong Kong: A Cross-Sectional Survey. *J. Med. Internet Res.* **2019**, *21*, e10831. [[CrossRef](#)]
42. Wu, S.; Cohen, D.; Shi, Y.; Pearson, M.; Sturm, R. Economic Analysis of Physical Activity Interventions. *Am. J. Prev. Med.* **2011**, *40*, 149–158. [[CrossRef](#)]
43. Zhu, H.; An, R. Impact of home-delivered meal programs on diet and nutrition among older adults: A review. *Nutr. Health* **2013**, *22*, 89–103. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.