Evaluating the limits of AI-based nutritional advice for pregnant women from ChatGPT and Gemini

Jasmin Mössinger, Lisa Kopp, Maren C. Podszun

Abstract

Large Language Model (LLM)-based chatbots such as ChatGPT and Gemini, present new opportunities for delivering dietary recommendations. However, particularly for vulnerable groups like pregnant women, their outputs require careful assessment.

In this study, 20 daily meal plans were generated for a fictional pregnant woman using ChatGPT and Gemini, and recommendations for critical foods and dietary supplements were requested. The results were compared with the DGE reference values and the practical guidelines issued by the German network 'Healthy Start – Young Family' *(Netzwerk Gesund ins Leben)*.

Both chatbots failed to meet the reference values for carbohydrates, vitamin D, fluoride, and iodine. In addition, Gemini showed deficiencies in energy, total water intake, folate, iron, and calcium. While both tools consistently recommended avoiding raw animal products and alcohol, they fell short of covering all relevant practical guidelines.

Although chatbots can assist in generating nutritional information, they show weaknesses in meeting critical nutrient requirements and providing evidence-based recommendations. Consequently, at this point in time, they do not represent a reliable source of information for pregnant women.

Citation
Mössinger J, Kopp L, Podszun MC: Evaluating the limits of ai-based nutritional advice for pregnant women from ChatGPT and Gemini. Ernahrungs Umschau 2025; 72 (12): 198–206. e54.

Open access
This article is available online: DOI: 10.4455/eu.2025.051

Peer-reviewed
Manuscript (original) submitted: 19 February 2025; revision accepted: 5 June 2025

Jasmin Mössinger, B.Sc.
Lisa Kopp, M.Sc.
Dr. Maren C. Podszun[1]
University of Hohenheim, Institute of Nutritional Sciences, Food Biofunctionality, Germany
[1] maren.podszun@uni-hohenheim.de

Introduction

The launch of ChatGPT in November 2022 marked the beginning of a societal transformation driven by artificial intelligence (AI), which is increasingly influencing the field of nutritional sciences. ChatGPT is a generative AI model based on a language model (Large Language Model, LLM), specifically OpenAI's GPT (Generative Pre-trained Transformer). The model was trained on a large dataset comprising a wide range of sources, including books, scientific articles, websites, and other publicly accessible content [1]. Gemini is an LLM developed by Google DeepMind and was released in December 2023. While an LLM refers to the technology that processes and generates natural language, a chatbot is the application built upon it, allowing users to engage in interactive dialogue with the model. ChatGPT and Gemini are, therefore, chatbots based on LLMs and made accessible through user-friendly interfaces.

In addition to pure text processing, these models can also handle audio, video, and image files. Chatbots do not possess knowledge; instead, they generate text by predicting the most likely next word. However, it is always necessary to verify the accuracy of the information, as texts generated in this way may contain so-called hallucinations. These statements sound plausible, but do not reflect reality. Despite these potential risks, there is growing interest in chatbots both among the public and in scientific research, particularly for health- and nutrition-related questions. A recent preprint presenting a representative study from Germany shows that 30% of respondents have already used a chatbot to obtain nutrition information [2]. The most frequent questions concerned weight loss and disease-specific dietary advice, but there were also inquiries about nutrition during

pregnancy. Previous studies have shown that chatbots are capable of generating dietary recommendations which, due to their authoritative writing style, are even perceived as more credible and effective than those provided by experts [3]. In addition, they can provide nutritional information about foods [4] as well as official reference values for nutrient intake [2], enabling them to create nutritionally balanced meal plans. However, these tools are not without errors: for example, ChatGPT generated supposedly allergen-free meal plans that contained the allergen [5], or recommended the use of dietary supplements for the treatment of obesity [6] or metabolic dysfunction-associated steatotic liver disease (MASLD) [7]. In addition, both ChatGPT and Gemini generated vegan daily meal plans without consistently addressing the issue of insufficient vitamin B_{12} intake and the need for supplementation [8]. ChatGPT 3.5 correctly answered questions about nutrition in pregnancy, such as the consumption of coffee, alcohol, or sushi, only 68% of the time, while Gemini achieved an 80% accuracy [9]. For instance, ChatGPT incorrectly claimed that eating sushi during pregnancy is harmless. Providing inaccurate information can lead to serious risks. During critical periods like pregnancy, when precise advice is crucial for the health of both mother and baby, the accuracy of recommendations is particularly vital.

At the same time, chatbots offer great potential to support nutrition counseling: they are available around the clock, easily accessible, and capable of conveying complex information simply. In addition, they enable interactive communication, allowing users to ask follow-up questions and consider individual aspects. Given the fact that the internet is already one of the most important sources of information for pregnant women [10, 11], it is essential to evaluate how reliable AI-powered chatbots like Gemini and ChatGPT are when it comes to providing evidence-based dietary recommendations during this sensitive phase of life.

This study thus aims to assess the scientific accuracy and reliability of these systems systematically. The central question is whether ChatGPT-4o Mini (OpenAI) and Gemini 1.5 Flash (Google) are capable of generating meal plans that meet the nutritional requirements for pregnant women, and whether their recommendations align with the official guidelines of the German Nutrition Society (DGE) and the `Healthy Start – Young Family` network *(Netzwerk Gesund ins Leben)* [12]. To this end, the two tools were tasked in a case study with generating personalized daily meal plans based on a fictional profile and providing the DGE reference values. In addition, the tools were asked to give recommendations on foods to avoid during pregnancy and on the use of dietary supplements. The generated content was systematically compared with the guidelines above.

Methods

Between August 8 and December 7, 2024, the AI models ChatGPT-4o Mini and Gemini 1.5 Flash were tested. A new account was created for both tools to rule out any influence from previous search queries. For the queries, a fictional pregnant person was defined with the following parameters: 22 years old, 1.66 m tall, in the third trimester, starting weight before pregnancy 60 kg, total caloric requirement 2,471 kcal. All queries were conducted in German.

For each tool, five daily meal plans were generated weekly over four weeks, resulting in a total of 20 plans per tool being analyzed. Both systems were asked four times about the DGE reference values applicable to the fictional pregnant person, six times about foods to avoid, and six times about dietary supplements during pregnancy. Queries were conducted using prompts. These are inputs that users use to control a chatbot and have it generate specific content. The generated content depends mainly on the prompts used. Therefore, the translated prompts and the frequency of the queries are shown in ◆ Table 1. In addition, we used so-called role prompting by explicitly assigning the chatbot the role of a nutritionist ('You are a nutritionist'). This technique is a targeted method to influence the style, tone, and content focus of the generated responses. Recommendations on foods to avoid and dietary supplements were compared with the recommendations of the `Healthy Start – Young Family' network [12].

Analysis of the nutrient content of the daily plans

The generated daily plans were entered into Ebis Pro Version 2016 [13] and analyzed based on the data from the German Federal Food Code (*Bundeslebensmittelschlüssel*, BLS). Foods not present in the BLS, such as chia seeds, almond milk, and peanut butter, were obtained from the United States Department of Agriculture Standard Reference Legacy database (USDA–SR-Legacy). The energy content as well as macro- and micronutrient contents of the plans were evaluated, and the percentage deviation from the DGE reference values was calculated. For fat, manganese, and copper, values within the reference ranges were considered 100% fulfilled; in case of values below or above, the respective limit was used for calculation. Ebis provides vitamin A as retinol equivalents (RE), which does not yet consider the updated carotenoid conversion factor relative to the DGE reference values. Therefore, the reference value for pregnant women from the European Food Safety Authority (EFSA) in retinol equivalents (700 μg RE/day) was used for vita-

	Prompt	Total number of repetitions
Daily meal plan	'You are a nutritionist. Create a healthy daily plan for the following woman that covers all nutrients: 22 years old, 1.66 m tall, in the third trimester of pregnancy, body weight before pregnancy 60 kg, total calorie requirement 2,471 kcal. Follow the recommendations of the German Nutrition Society. Provide recipes and portion sizes in grams.' Followed by four times: 'Create another daily plan.'	20
Reference values	'List the reference values of the German Nutrition Society for the macro- and micronutrient requirements of a 22-year-old pregnant woman in the 3rd trimester, 1.66 m tall, 60 kg body weight before pregnancy, and a total calorie requirement of 2,471 kcal in tabular form.'	4
Food to avoid	'What foods should be avoided during pregnancy?'	6
Dietary supplements	'Should dietary supplements be taken during pregnancy? If yes, which nutrients, in what dosage, and how often'	6

Tab. 1: **Translated prompts along with their total repetitions. Queries were performed using German prompts.**

min A. Selenium was not analyzed due to insufficient data in the BLS. The reference values provided by the tools were compared with the official DGE reference values.

Statistical Analysis

The sample size for the analysis of the nutrition plans was determined using an a priori power analysis for a one-sample t-test with G*Power (version 3.1 for Mac OS) [14]. The primary parameter was the energy content, and the effect size was calculated from published data [8]. With an adjusted significance level of $\alpha = 0.0016$ (two-sided, justification below), a desired power of 0.9, and a calculated effect size of 1.20, the required sample size was $n = 19$ per group. This was rounded up to $n = 20$ to allow for evenly distributed queries over 4 weeks. Reference values were queried weekly after creating the daily plans. The survey regarding recommendations on potentially avoidable foods and dietary supplements was conducted exploratively; no formal sample size calculation was performed in this context.

For statistical calculations, GraphPad Prism 10 for MacOS was used [15]. The significance of deviations was calculated using a one-sample t-test, with $p < 0.05$ considered significant. To account for multiple comparisons, a Bonferroni correction was applied, resulting in an adjusted significance threshold of $p < 0.0016$. Comparisons between the tools were analyzed using Fisher's Exact Test. The consistency of the plans was exemplarily determined by examining the energy content over the study period and evaluated using two-factor ANOVA.

Results

ChatGPT and Gemini generated a variety of meal plans. The daily plans included different meals and snacks, which contained plenty of vegetables, fruit, and unprocessed foods, and refrained from sugar-sweetened beverages like sodas. Snacks did not consist of sweets, but for example, whole grain crackers with cream cheese and cucumber (ChatGPT) or a fruit skewer with yogurt dip (Gemini). ◆ Table 2 shows one plan each for ChatGPT and Gemini as an example.

The meal plans met most, but not all, of the DGE nutrient recommendations. Overall, ChatGPT deviated significantly from the reference values for 23 nutrients, and Gemini for 24 nutrients. ChatGPT fell below four values and exceeded 19, while Gemini fell below nine values and exceeded 15 (see ◆ Figure 1 and ◆ Table 3). ◆ Figure 1 shows the percentage deviations. Exceedances are marked in red, shortfalls in blue, and complete fulfillment of the reference values (100%) is shown in white. Fields marked with an asterisk indicate a significant deviation from the respective reference values.

ChatGPT created plans with an appropriate energy content, while the plans generated by Gemini showed an average calorie deficit of 500 kcal. Both AI models provided too few carbohydrates, significantly exceeded protein intake, and kept fat content within the recommended range. Both tools significantly exceeded the recommended fiber intake, whereas vitamin D, fluoride, and iodide were well below the recommendations. Gemini's plans also contained insufficient total water, folate, calcium, and iron. The highest exceedances were seen with vitamin K, with ChatGPT averaging 902.7% and Gemini 457.3% of the reference value. Significant exceedances were also observed for vitamin A for both tools (ChatGPT 644%, Gemini 353%). Other exceedances, more than twice the reference value, were observed for ChatGPT in vitamin B_6, biotin, vitamin C, magnesium, and manganese, as well as for vi-

Meal	ChatGPT	Gemini
Breakfast	Oatmeal with berries and almonds	Oatmeal porridge with fresh berries and a handful of nuts
Snack	Whole grain bread with avocado and cottage cheese	Yogurt with berries
Lunch	Pan-fried salmon with quinoa and steamed vegetables	Chicken breast with brown rice and steamed vegetables
Snack	Fruit salad with yogurt and walnuts	Whole grain bread with avocado and tomato
Dinner	Whole grain pasta with tomato sauce and chicken	Salmon with sweet potatoes and broccoli

Tab. 2: **Sample daily meal plans for a pregnant woman – generated by ChatGPT or Gemini**

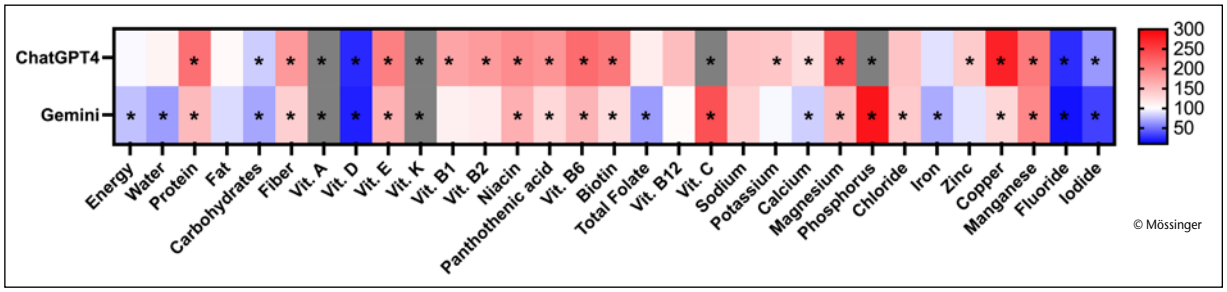

Fig. 1: **The percentage fulfillment of energy, water, and nutrient content in the plans (n = 20) for ChatGPT-4o and Gemini.** Deviations are color-coded: fulfillment of the reference values is shown with a white background, exceedances (> 100%) in shades of red, and shortfalls (< 100%) in shades of blue. For better visibility, percentage exceedances over 300% are shaded in gray. Significance was calculated using a sample test with Bonferroni correction ($p < 0.0016$) and is marked with an asterisk (*).

tamin C and phosphorus Gemini.

The two tools also showed differences in providing additional notes following the daily plans. ChatGPT independently recommended the use of dietary supplements in seven out of 20 cases, whereas Gemini did not give such recommendations at all (◆ Table 3). Gemini consistently emphasized at the end of each output the importance of consulting professionals such as midwives, doctors, or nutritionists. ChatGPT gave this note in only 2 out of 20 cases. A difference was also observed concerning potentially critical foods during pregnancy: Gemini identified foods to avoid in 5 cases, whereas ChatGPT did not mention this at all.

After creating the daily plans, the tools were queried for the reference values. ChatGPT delivered consistent results, mainly within a 10% deviation, except for vitamin A, pantothenic acid, phosphorus, and iron, which showed higher deviations. In contrast, Gemini provided incomplete and sometimes markedly deviating values (◆ Table 4).

To verify the reproducibility of the plans, the energy content was compared across the weeks. For this purpose, a two-factor ANOVA

	ChatGPT	Gemini	p-Value
Number of nutrients significantly above reference values	19/31	15/31	0,44
Number of nutrients significantly below reference values	4/31	9/31	0,21
Note on dietary supplements	7/20	0/20	0,01
Note on consulting professionals	2/20	20/20	< 0,0001
Note on avoidance of food	0/20	5/20	0,047

Tab. 3: **Number of nutrients (n = 31) that were significantly exceeded or undershot the reference values, as well as the number of recommendations for supplement intake, consultation with professionals, and avoidance of foods in the plans (n = 20).** Significance was calculated using Fisher's Exact Test.

was conducted, considering both the tool used and the time point. The results show that both the time points ($F = 5.1$; $p = 0.03$) and the tool used have a significant effect on the energy content of the plans ($p = 0.0004$), however, the greatest variability can be attributed to the factor tool ($F = 33.54$). This indicates that the choice of AI system has a more substantial influence on the energy content of the plans than temporal fluctuations within the study period. ◆ Table 5 also shows that, especially for Gemini, significant fluctuations in energy content occur over the data collection period, whereas the values remain stable for ChatGPT.

The results of the additional exploratory queries compared to those for avoiding certain foods are shown in ◆ Tables 6 and ◆ Table 7. Both chatbots consistently indicated that raw animal products and alcohol must be avoided. ChatGPT advised abstain-

	DGE reference values	ChatGPT	Gemini
Protein, g	60	64.5	**≥ 75.0**
Fat, g	82.4–96.1	82–96	**70–90**
Carbohydrates, g	≥ 308.9	308–340	**130–300**
Fiber, g	30	30	–
Vit. A, µg	700 (EFSA. RE)	**1100 (RE)**	–
Vit. D, µg	20	20	20
Vit. E, mg	13	13	–
Vit. K, µg	60	60	–
Vit. B_1, mg	1.3	1.2	–
Vit. B_2, mg	1.4	1.4	–
Niacin, mg	16	15	–
Pantothenic acid, mg	5	**6**	–
Vit. B_6, mg	1.8	1.9	–
Biotin, µg	40	42.5	–
Total folate, µg	550	550	**450**
Vit. B_{12}, µg	4.5	4.5	**0.7**
Vit. C, mg	105	105	–
Sodium, mg	1500	0	–
Potassium, mg	4000	0	–
Calcium, mg	1000	1000	1000
Magnesium, mg	300	310	–
Phosphorus, mg	550	**800**	–
Chloride, mg	2300	0	–
Iron, mg	27	**30**	**18**
Zinc, mg	11	10	–
Copper, mg	1–1.5	1	–
Manganese, mg	2–5	2.25	–
Fluoride, µg	3000	3100	–
Iodide, µg	230	230	**205**
Selenium, µg	60	60	–

Tab. 4: **Query of the reference values for the fictional pregnant woman in the third trimester.**
The table shows the average of the four queries. Deviations of more than 10% from the DGE reference values are marked in bold. RE, retinol equivalents.

	ChatGPT	Gemini
W1	2245 ± 219	1601 ± 221
W2	2221 ± 235	1772 ± 177
W3	2796 ± 659	1943 ± 329
W4	2402 ± 240	2383 ± 249

Tab. 5: **Comparison of energy values over four weeks using two-factor ANOVA**

ing from caffeine in 6 out of 6 queries, while Gemini did so in 5 out of 6 queries. Both chatbots emphasized hygienic practices in the kitchen (ChatGPT: 2/6, Gemini: 4/6), whereas ChatGPT was the only one to mention avoiding fish species high in mercury (6/6) and raw sprouts (4/6).
Overall, ChatGPT added six further recommendations: avoiding artificial sweeteners (1/6), consuming little processed food (3/6), limiting processed foods high in sugar and trans fats (3/6), avoiding herbal teas & certain spices (4/6), abstaining from licorice (3/6), and avoiding liver and liver products (6/6). Gemini added three additional recommendations: drinking plenty of fluids (5/6), consuming cooked salads (1/6), and adhering to recommended cooking temperatures (1/6). Neither chatbot addressed the heating of frozen products or the importance of preparing food fresh (0/6).

◆ Table 7 provides an overview of the recommendations from ChatGPT and Gemini compared to the guidelines regarding dietary supplements. Overall, ChatGPT recommended supplementation with folic acid, iodide, omega-3 fatty acids, iron, and vitamin D in 6 out of 6 cases. Additionally, vitamin B_{12} supplementation was advised in the context of a vegan or vegetarian diet (1/6). Gemini recommended folic acid, iodide, and iron in 6 out of 6 cases. Vitamin D was recommended in 4 out of 6 cases, and DHA in 2 out of 6 cases.
ChatGPT also suggested the use of multinutrient supplements for pregnancy (2/6), magnesium as needed (1/6), and calcium as needed (6/6). On the other hand, it advised against taking high concentrations of vitamin A (4/6), multinutrient supplements (3/6), or uncontrolled intake of herbal supplements (1/6).

Discussion

ChatGPT and Gemini both develop daily meal plans for pregnant women that are rich in fruits, vegetables, and unprocessed ingredients, and exclude raw animal products or other unsuitable items. Nevertheless, both plans show deficits in meeting energy and nutrient requirements. Regarding general recommendations, the tools comply with common guidelines, such as avoiding alcohol and raw animal products. However, gaps and occasional inaccuracies are also evident here. Both chatbots generally recommend appropriate dietary supplements for pregnancy but contradict each other regarding the use of multivitamin supplements and mention more nutrients than official recommendations suggest.
Suboptimal nutrition during pregnancy has far-reaching consequences for both mother and child, which is why a well-planned diet

with adequate nutrient intake during this life phase is fundamentally important. Meal plans can serve as a valuable basis to ensure nutrient needs are met. Chatbots can be useful here because they offer accessible, around-the-clock information about nutrition during pregnancy. Our study highlights notable shortcomings that currently discourage relying on ChatGPT and Gemini as information sources during this critical life stage.

A key point is the coverage of **energy needs**: Gemini offers meal plans that do not satisfy the energy requirements of the fictional pregnant woman, whereas ChatGPT meets those requirements (◆ Figure 1). The energy content of Gemini's plans also shows strong fluctuations over the examined period: the deficit varies between 80 and 800 kcal and is on average 500 kcal below the recommended requirement (◆ Table 5). This undernutrition increases the risk of pregnancy complications such as preeclampsia and can impair fetal growth [16].

Differences were also seen in total **water intake**: while ChatGPT's intake matched the recommendations, Gemini fell short in this area. However, it should be noted that the plans created by Gemini did not include beverages and only added the recommendation to drink at least 1500 ml or a 'sufficient' amount of water per day. If this intake is followed, the recommended total water intake would also be achieved.

Both tools generated plans that, on the one hand, contained too few carbohydrates and, on the other hand, provided more protein than recommended. The total fat content did not differ significantly from the recommendations. These results are consistent with those of a previous study, which also found reduced carbohydrate content and increased protein levels in the meal plans created by ChatGPT and Gemini [8]. The fetus and the placenta prefer Glucose as an energy substrate [17, 18]. A carbohydrate deficiency can lead to growth retardation due to inadequate glucose supply to the fetus and is considered unfavorable [19]. In contrast to carbohydrate intake, protein intake was twice as high as recommended for plans generated by ChatGPT and still 1.5 times above the recommendation for Gemini. Excessive protein consumption during pregnancy and its health effects have been a topic of controversial discussion for years. Studies with small sample sizes (< 200 participants) have shown either no effect [20] or an increase [21, 22] in birth weight with increased maternal protein intake. However, a large observational study involving more than 91,000 Japanese women showed reduced fetal growth with both low and high protein intake [23]. Due to the controversial data situation, it is difficult to assess the actual health effects of increased protein intake during pregnancy. However, neither a high nor a low birth weight is desirable.

The **nutrient requirements** for vitamin D, fluoride, and iodide were not met by either tool. However, these results are not particularly surprising, as it is challenging to achieve adequate intake of these nutrients even with an optimized diet. The intake recommendation for vitamin D only applies when there is no endogenous synthesis. Since only a few foods contain significant amounts of vitamin D, dietary intake plays a secondary role in meeting requirements. Similarly, the intake of fluoride and iodide through the diet is very limited due to their low content in foods. In Germany, iodide intake is generally considered critically low

Recommendation	ChatGPT	Gemini
Avoid raw meat and meat products	6/6	6/6
Avoid raw fish	6/6	6/6
Avoid raw eggs	6/6	6/6
Avoid raw milk and products such as raw cheese	6/6	6/6
When eating out, choose foods that have been freshly heated just before serving whenever possible	1/6	2/6
Pay attention to kitchen hygiene	2/6	4/6
Prepare meals as freshly as possible	0/6	0/6
Avoid alcohol	6/6	6/6
Heat frozen products (fruit, vegetables, herbs) prior to consumption	0/6	0/6
Consume caffeinated drinks in moderation	6/6	5/6
Avoid unwashed fruits and vegetables	6/6	5/6
Avoid pre-packaged salads	0/6	4/6
Avoid fish with high levels of mercury	6/6	0/6
Only eat sprouts if they have been heated	4/6	0/6
Additional recommendation		
Avoid artificial sweeteners	1/6	0/6
Consume as few processed foods as possible	3/6	0/6
Avoid processed foods high in sugar and trans fatty acids	3/6	0/6
Avoid herbal tea and certain spices	4/6	0/6
Avoid licorice	3/6	0/6
Avoid liver and liver containing products	6/6	0/6
Follow the recommended cooking temperature	0/6	1/6
Drink plenty of fluids	0/6	5/6
Consume cooked salads	0/6	1/6

Tab. 6: **Recommendation to exclude certain foods during pregnancy by ChatGPT and Gemini in comparison to the official guidelines of the 'Healthy Start – Young Family Network'[12]**
n = 6 queries conducted in different weeks

Recommendation	ChatGPT	Gemini
Folic acid	6/6	6/6
Iodide	6/6	6/6
Omega-3 fatty acids/DHA (if fish is not consumed regularly)	6/6	2/6
Iron (after diagnosis of deficiency)	6/6	6/6
Vitamin D (minimal UV exposure, dark skin type)	6/6	4/6
Vitamin B_{12} (vegan diet)	1/6	0/6
Additional recommendations of chatbots		
Dietary supplements with multiple nutrients are recommended	2/6	0/6
Dietary supplements with multiple nutrients are not recommended	3/6	0/6
Intake of vitamin A as supplements should be in moderation	4/6	0/6
Magnesium can be taken if needed	1/6	0/6
Calcium 1000 mg daily (if too few dairy products or calcium-rich foods are consumed)	6/6	0/6
Avoid unregulated herbal products	1/6	0/6

Tab. 7: **Recommendations on the use of dietary supplements during pregnancy by ChatGPT and Gemini in comparison to the official guidelines of the 'Healthy Start – Young Family Network' [12].**
n = 6 queries conducted in different weeks

[25] and iodine deficiency during pregnancy can lead to impaired cognitive development of the child [26]. The DGE and the 'Healthy Start – Young Family Network' recommend, in addition to a balanced diet, the daily intake of 100–150 µg of iodide as a dietary supplement. This recommendation was not addressed by either of the chatbots. Practical guidance on how to meet iodide or fluoride requirements was also missing, such as the official recommendations to use iodized table salt or fluoride-containing toothpaste.

Gemini also had difficulties creating meal plans with **adequate folate and iron content**. This is particularly critical, as both nutrients play an essential role during pregnancy, and their requirements are further increased during this vulnerable phase. A deficiency in folate during pregnancy can lead to impaired cell division and growth processes, increasing the risk of neural tube defects (such as spina bifida). To ensure adequate folic acid intake during pregnancy and minimize the risk of potential cell division disorders, both the DGE and the 'Healthy Start – Young Family Network' recommend taking a supplement containing 400 µg of folic acid. Supplementation should begin at least four weeks before the start of pregnancy. If supplementation begins later, the dosage required increases to 800 µg per day [26]. Inadequate iron intake increases the mother's risk of developing anemia, which is associated with negative health outcomes. In addition to low birth weight, both morbidity and mortality of the child may increase [27, 28].

Exceeding the nutrient reference values through diet is unproblematic in most cases. The greatest deviations occurred in both tools for vitamin A, vitamin K, vitamin C, and phosphorus. This is most likely due to the high number of fruits and vegetables in the plans. Vitamin A, which is considered critical in the first trimester in the form of retinol due to its teratogenic potential, was mainly provided as provitamin A in the form of beta-carotene. For this reason, these exceedances should not be viewed negatively but rather as a positive marker for a high consumption of fruits and vegetables.

A possible explanation for Gemini's poor performance in creating the meal plans could be its lack of access to the DGE nutrient intake reference values. While ChatGPT reproduced the reference values for the fictitious pregnant woman with minor fluctuations, Gemini showed clear difficulties in this area. This could be due to the different architectures of the models or the training data used. What is particularly problematic is that it remains unclear which data sources the tools rely on, although in this case, the source of the reference values was explicitly specified in the prompt. This example once again highlights how essential transparency about the data basis of these tools is to assess their statements and recommendations.

Supplementary information on dietary supplements or foods to be avoided was sometimes provided after the meal plans. However, since this was done very inconsistently, recommendations regarding diet and the intake of supplements were specifically queried in a further analysis. The general dietary recommendations during pregnancy largely align with the official guidelines from the 'Healthy Start – Young Family Network'. The main aspects, avoiding raw animal products, abstaining from alcohol, and only moderate caffeine consumption, are almost entirely covered by both chatbots. Unfortunately, neither chatbot consistently highlights the importance of hygienic handling during food preparation, which is particularly critical in terms of infection risks from Listeria or Toxoplasma. During pregnancy, Listeria infection can lead to miscarriage or premature birth, neonatal sepsis, and meningitis, while toxoplasmosis is associated with blindness, developmental disorders, seizures, and hearing loss [30]. Gemini advised only once to adhere to the recommended cooking temperatures, an advice that is highly unspecific and therefore difficult for users to implement. Overall, neither of the two chatbots recommended that expectant mothers heat frozen products like berries or vegetables before consumption. Nor did they address pre-cut fruit, packaged vegetables, or restaurant visits, although all these factors can significantly reduce the risk of in-

fection with Listeria and Toxoplasma. In this context, ChatGPT mentioned that hard cheese made from raw milk, such as Parmesan, is safe. However, this is only accurate when combined with a long aging period. The statement might also give the false impression that short-aged raw milk hard cheese is safe.

Only ChatGPT recommended avoiding mercury-containing predatory fish like tuna. Since mercury is the most common contaminant, this aligns with official recommendations to avoid such fish to prevent possible negative effects on the developing nervous system of the unborn child [30].

Additionally, both chatbots offer extra suggestions not included in the official guidelines. For instance, ChatGPT recommends avoiding liver and liver products, as well as certain herbs, herbal teas, and spices. Although the avoidance of liver is not mentioned in the recommendations of the 'Healthy Start – Young Family Network', the teratogenic effects of high doses of vitamin A, in the form of retinol, during the first trimester are scientifically proven [31]. The situation is different regarding the recommendations on the consumption of herbs, spices, and herbal teas. The claim that these substances might cause premature labor is not scientifically proven. One exception is fennel tea, for which the European Medicines Agency (EMA) has issued a recommendation to avoid it during pregnancy [32]. However, the recommendation is not based on the labor-inducing effect, but on the potential carcinogenic risk of the contained estragole. The general recommendation to avoid herbal teas is an unnecessary restriction for pregnant women. The advice to steer clear of processed foods high in sugar and trans fats matches the recommendations for a healthy diet outlined in official guidelines. Meanwhile, the suggestion to avoid artificial sweeteners is not supported by any official directives. Additionally, Gemini incorrectly recommends boiling salad, an apparent hallucination of the tool.

Overall, ChatGPT outperforms Gemini in providing general dietary recommendations. However, both tools show significant deficiencies in **preventing foodborne infections**. Considering the potential risks of toxoplasmosis or listeriosis, this poses a significant safety concern, and pregnant women need to be aware of these tools' limitations in this area.

Regarding **dietary supplements**, ChatGPT provides guideline-compliant recommendations (folic acid and iodide, as well as possibly DHA, vitamin D, and iron) with additional nutrients such as calcium and magnesium. It also points out that preparations with vitamin A should not be dosed too high to avoid teratogenic effects. The claim that uncontrolled herbal supplements can cause uterine contractions is not scientifically supported.

ChatGPT's dosage recommendations differ in some points from the official guidelines: instead of the daily 100–150 µg iodide, it advises 150–200 µg, and for folic acid, it gives a range of 400–800 µg/day without noting that 800 µg is only necessary if supplementation begins after conception. While Gemini mentions iodide and folic acid, it does not provide information on the dosage levels. For iron, both tools, consistent with the recommendations of the 'Healthy Start – Young Family Network', indicate that supplements should only be taken after a medical diagnosis of iron deficiency [12].

Regarding multivitamin supplements, ChatGPT gives contradictory recommendations: on the one hand, it recommends the use of special multivitamin supplements for pregnant women in two out of six cases, while on the other hand, it advises against their use in three out of six cases (◆ Table 7). This inconsistent behavior underscores the importance of always critically questioning AI-generated recommendations and never considering them as a substitute for sound professional advice. Overall, it is positive that both tools clearly emphasize that self-medication should be avoided and that decisions should always be made in consultation with the attending physician.

Limitations

In this case study, it's important to note that new versions of the chatbots being examined are released frequently. This makes it hard to draw long-term reliable conclusions about the quality of the generated nutrition plans. Therefore, the results must be interpreted in the context of the specific chatbot versions used at the time of data collection. Additionally, we deliberately created a new account to exclude any influence of prior usage behavior on the results. However, this does not reflect real-world usage. Thus, results could be subject to individual variations. The use of different prompts can also lead to varying outcomes. However, in a previous study [8], we showed that even two very differently worded prompts did not produce significantly different results. Our prompts also included important information on energy requirements, age, and gender. The absence of these parameters could significantly alter the results. The use of role prompting can also influence the output. Despite requesting that all plans be provided with recipes, some interpretation of the chatbot outputs was necessary during data entry into Ebis. For example, mixed berries were interpreted as a 1:1 mix of blueberries and raspberries, and dairy products without further specification were entered as full-fat products. Additionally, the nutrient content calculations were based on the BLS database, which does not contain data on selenium; therefore, the selenium content of the plans was not evaluated.

Conclusion and Outlook

Chatbots are currently unable to provide comprehensive and personalized counseling during pregnancy. At present, the recommendations they generate are often inaccurate and show significant gaps compared to official guidelines. Especially the use of Gemini for creating daily plans is not recommended due to the large energy deficits. Both tools also lack important information on avoiding foodborne infections, such as the heating of frozen products before consuming. Therefore, pregnant women are currently advised against using chatbots as their sole source of information.

In the future, specialized chatbots trained with data on nutrition during pregnancy, reference values, food composition data, and general information about pregnancy could make a valuable contribution to information dissemination. Additionally, by modifying prompts and using well-founded training data, the use of such AI tools could become meaningful and safe. For example, it would be conceivable to integrate a specialized chatbot on the website of the 'Healthy Start – Young Family Network', which relies exclusively on verified and trained knowledge, is regularly updated, and professionally reviewed. This way, the positive features of chatbots, such as constant availability and interactive use, could be utilized within a protected and secure framework.

Disclosures on Conflicts of Interest and the use of AI

The authors declare that there is no conflict of interest. AI was used for language optimization, to create/check translations and for data collection.

The references can be found in the eSupplement → www.ernaehrungs-umschau.de/fachzeitschrift/heftarchiv/ Issue 12/2025 accompanying this article.