

Article

Precision Nutrient Management Using Artificial Intelligence Based on Digital Data Collection Framework

Hsiu-An Lee ¹, Tzu-Ting Huang ², Lo-Hsien Yen ², Pin-Hua Wu ³, Kuan-Wen Chen ¹, Hsin-Hua Kung ¹,
Chen-Yi Liu ^{2,*} and Chien-Yeh Hsu ^{2,4,*}

- ¹ National Health Research Institutes-The National Institute of Cancer Research, Tainan 704, Taiwan; hsiuanlee@nhri.edu.tw (H.-A.L.); hank9134277698@nhri.edu.tw (K.-W.C.); hsinhuakung@nhri.edu.tw (H.-H.K.)
- ² Department of Information Management, National Taipei University of Nursing and Health Sciences, Taipei 112, Taiwan; 102216003@ntunhs.edu.tw (T.-T.H.); 102216013@ntunhs.edu.tw (L.-H.Y.)
- ³ College of Health Technology, National Taipei University of Nursing and Health Sciences, Taipei 112, Taiwan; 102016002@ntunhs.edu.tw
- ⁴ Master Program in Global Health and Development, College of Public Health, Taipei Medical University, Taipei 110, Taiwan
- * Correspondence: 102216016@ntunhs.edu.tw (C.-Y.L.); cyhsu@ntunhs.edu.tw (C.-Y.H.)

Abstract: (1) Background: Nutritional intake is fundamental to human growth and health, and the intake of different types of nutrients and micronutrients can affect health. The content of the diet affects the occurrence of disease, with the incidence of many diseases increasing each year while the age group at which they occur is gradually decreasing. (2) Methods: An artificial intelligence model for precision nutritional analysis allows the user to enter the name and serving size of a dish to assess a total of 24 nutrients. A total of two AI models, including semantic and nutritional analysis models, were integrated into the Precision Nutritional Analysis. A total of five different algorithms were used to identify the most similar recipes and to determine differences in text using cosine similarity. (3) Results: This study developed two models to form a precision nutrient analysis model. The 2013–2016 Taiwan National Nutrition Health Status Change Survey (NNHS) was used for model verification. The model's accuracy was determined by comparing the results of the model with the NNHS. The results show that the AI model has very little error and can significantly improve the efficiency of the analysis. (4) Conclusions: This study proposed an Intelligence Precision Nutrient Analysis Model based on a digital data collection framework, where the nutrient intake was analyzed by entering dietary recall data. The AI model can be used as a reference for nutrition surveys and personal nutrition analysis.

Keywords: nutrition survey; precision diet analysis; medical intelligence



Citation: Lee, H.-A.; Huang, T.-T.; Yen, L.-H.; Wu, P.-H.; Chen, K.-W.; Kung, H.-H.; Liu, C.-Y.; Hsu, C.-Y. Precision Nutrient Management Using Artificial Intelligence Based on Digital Data Collection Framework. *Appl. Sci.* **2022**, *12*, 4167. <https://doi.org/10.3390/app12094167>

Academic Editors: Neil Yuwen Yen, Chen-Kun Tsung, Vinod Kumar Verma and Chao-Tung Yang

Received: 17 February 2022

Accepted: 19 April 2022

Published: 20 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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

Nutritional intake is the basis for human growth and health, and the intake of different types of nutrients and micronutrients can affect health. Most diseases are inextricably linked to diet. Diabetes, cardiovascular diseases (hypertension, hyperlipidemia), gout, peptic ulcers, and gastroenteritis are all diet-related diseases that are increasing in prevalence every year, while the age group of those suffering from these diseases is gradually decreasing. The development of the Internet has made it possible to conduct online nutrition surveys through large-scale food and nutrition databases linked to automated dietary records, and there are now a growing number of software, platforms, and applications for nutrition surveys [1].

The most common technologies used for dietary recording are web-based or online tools, mobile apps, camera-based image analysis tools, wearable sensors, etc., while traditional methods rely on the use of Food Frequency Questionnaires (FFQs) or 24 h dietary

recording methods. However, past techniques have suffered from a lack of accuracy in recording, as recall methods may not accurately record the food consumed or have difficulty estimating portion sizes or limited food ingredient lists [2].

The coding and translation of food records from nutrition surveys into nutrient analyses are labor-intensive and time-consuming, meaning that it is more difficult to collect detailed information regarding food intake in large scale population studies. Such studies rely on answers to food frequency questionnaires, and the accuracy of this data is dependent on the expertise of the interviewer compared to other self-reported measures [3,4].

Innovative technological tools have evolved with the development of various IT technologies, including natural language analysis of text, speech analysis, and image processing. The popularity of smartphones, tablets, and computers has increased the acceptance of using IT for nutritional intake assessments [5–8].

This study develops an artificial intelligence model for a precision nutrient analysis, which allows users to enter the name of a dish and serving size to assess a total of 24 nutrients. The recipes can be modified by the user, which allows the model to be used in all countries and all contexts, thus improving interoperability and accuracy of the analysis.

2. Related Works

The Food Record, the 24 h dietary recall (24HR), and the Food Frequency Questionnaires (FFQs) are three common methods of collecting nutritional data. The Food Record is a comprehensive record of all foods, beverages, and nutritional supplements consumed by the respondent over a specified period of time. Usually, 3–4 days of intake are recorded, as the quality (accuracy) of the record is reduced due to the burden of recording too many days. Ideally, dietary intake should be weighed and measured; however, most respondents only recorded pre and postestimates of intake, which would lead to differences in weight judgments [9].

The 24HR method assesses the nutritional intake of a respondent over the past 24 h. Ideally, the survey collects information on nutritional intake over multiple 24 h periods on nonconsecutive, random dates. The 24HR method is usually conducted by a dedicated interviewer by telephone or in person [9]. Some 24HR surveys can also be self-recorded or collected online (e.g., Automated Self-Administered 24 h dietary recall and ASA24 [10]). The differences between the ASA24 and 24HR methods primarily reduce interviewer burden and interview costs and allow respondents to answer questions at their own pace; however, this method may not be suitable for all study populations.

The use of exploratory questions in the 24HR recall method facilitates easy response and has been shown to improve the accuracy of data collection [11]. The survey includes how the food was prepared, what was added after preparation (seasonings, creams, and spices), and when the meal was served [9]. The FFQ assesses general nutritional intake over a specific period of time, usually a longer period, and asks how often a person consumes food. The FFQ method is a more cost-effective alternative to the 24HR method because respondents can complete the survey themselves, and it can be used for large sample studies [12].

There are several types of systematic measures of self-reported dietary information; for example, based on general perceptions, most respondents tend to report foods that are perceived as healthy and to report less on less healthy foods. However, differences in susceptibility to this tendency between groups of respondents can lead to additional personal bias. Differences in the ability to self-assess and recall portions can also lead to individual subjective differences. This systematic error is unpredictable, but studies suggest that it may be related to factors such as age and gender [13]. While each person uses different strategies to recall portion sizes, including taking photographs and using measurement aids to estimate (e.g., food models) [14,15], research shows that training can lead to a more accurate assessment of food portions [16,17]. In addition, researchers or the methods used to collect dietary data may also be biased [18].

Finally, the accuracy of the conversion of nutrient totals from nutrition dietary records depends on the accuracy and availability of the food ingredient database for conversion to calories and nutrients. In summary, both types of errors reduce the judgement of the relationship between diet and health, as well as the accuracy of the statistical analysis. However, while there may be some slight deviations in the database of the relationships tested [19], when the results of significant analyses are properly evaluated, valid conclusions can be drawn.

3. Materials and Methods

This study developed an AI model based on semantic text to analyze the nutritional ingredients of a nutrient, and a digital data semantic analysis model was designed to determine the names and servings of the dishes consumed. The AI model is based on the ingredients of common Taiwanese recipes and automatically calculates the nutrient intake. The model structure consists of a digital data semantic analysis model, an AI precision nutrient analysis model, a database of 1590 recipes, and 7869 ingredients from common Taiwanese recipe databases, and the model structure is shown in Figure 1. The nutrition information of the ingredients was obtained from the public data of the Health Promotion Administration, Ministry of Health and Welfare Taiwan (HPA, MoHW).

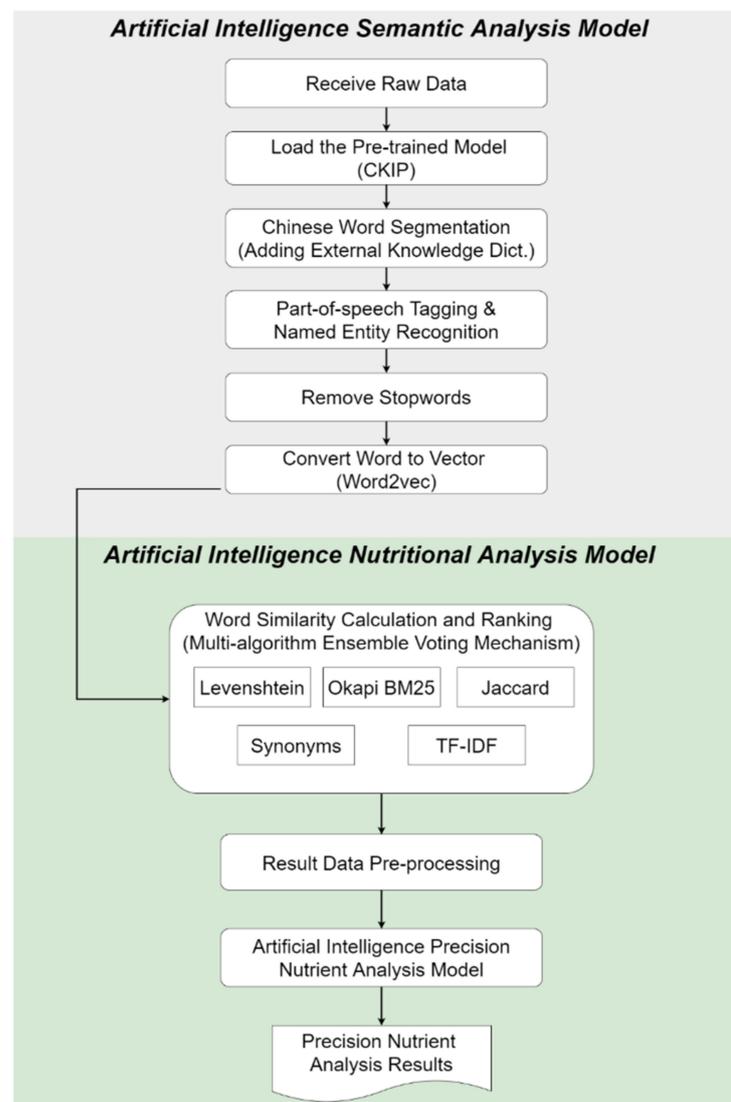


Figure 1. Model Structure.

3.1. Artificial Intelligence Semantic Analysis Model

Data were intercepted and annotated after data entry, and a CKIP pretraining model was used to interpret Chinese words. After completion, lexical annotation and entity identification were performed. Finally, the nouns (dish names) were converted into vector structures using word2vec, which is an application of Natural Language Processing proposed by Tomas Mikolov et al. at Google in 2013 and is one of the most significant advances in the field of machine learning in recent years. Word2vec is an application framework that learns large amounts of textual data and transforms words into mathematical vectors to discriminate their semantic meanings by embedding words into a two-dimensional space in order that words with similar semantic meanings can be closer together.

This study used the continuous bag-of-words (CBOW) method, which aims to determine the lexical properties of the input words using a whole paragraph of context and to determine the relationship between similar words by concatenating them. As similar words are clustered together, the direction of the vector corresponds to the relative relationship.

3.2. Artificial Intelligence Nutritional Analysis Model

The Nutritional Analysis Model is divided into three steps. Step 1 conducts artificial intelligence analysis to determine the most similar recipes. Due to the multicharacter nature of Chinese, single algorithm of semantic analysis may not be precise enough. Therefore, a variety of algorithms were used for the analysis. The AI model is composed of five different algorithms, including 1. Okapi BM25, 2. TF-IDF, 3. Levenshtein, 4. Jaccard, and 5. Synonyms. The algorithm also uses cosine similarity to determine differences in text and then compares it with a database to obtain food information and portion sizes for recipes and ingredient judgement. Step 2 is to determine the best solution by the common voting mechanism. Step 3 is nutritional ingredient calculation.

3.2.1. Step 1. Artificial Intelligence Analysis

(1) Okapi Best Matching (Okapi BM25)

This algorithm was proposed by Stephen E. Robertson, Karen Spärck Jones, and other scholars in 1970 [20–22]. As a probabilistic search framework, BM25 is still widely regarded as one of the most advanced ranking algorithms. BM25 is a bag-of-words model, which ranks a set of documents based on their similarity to each other and obtains a set of scores that can be compared with each other.

The BM25 similarity formula is shown in Equation (1).

$$\text{score}(D, Q) = \sum_{i=1}^n \text{IDF}(q_i) \left[\frac{f(q_i, D)(k_1 + 1)}{f(q_i, D) + k_1 \left(1 - b + b \frac{|D|}{\text{avgdl}}\right)} + \delta \right] \quad (1)$$

Equation (1) BM25 Similarity Formula

- $f(q_i, D)$: Frequency of the term q_i in Document D
- $|D|$: Length of Document D (in words).
- K_1 : The terminology described above is saturated with parameters.
- b : The length normalization parameters, as described above.
- avgdl : Average document length in document collection.
- IDF : Frequency of inverse text files.
- $n(q_i)$: Number of documents containing q_i .
- n : Total number of text files in the collection.

(2) Term Frequency–Inverse Document Frequency (TF–IDF)

This algorithm is a weighting technique widely used in information retrieval and text mining, and the combination of TF and IDF was first discussed by Karen Spärck Jones [23]. The TF–IDF was used to assess the importance of a word in a document, which increased positively with the number of times the word appears in the document but decreased

inversely with the frequency of its occurrence. The TF-IDF formula is shown in Equation (2), while the Inverse Document Frequency IDF Formula is shown in Equation (3).

$$tf_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}} \tag{2}$$

Equation (2) TF-IDF formula

- Molecular formula: $n_{i,j}$ denotes the number of occurrences of the word in document d_j .
- Denominator: The sum of all occurrences of the word in document d_j .

$$idf_i = \lg \frac{|D|}{|\{j : t_i \in d_j\}|} \tag{3}$$

Equation (3) Inverse Document Frequency IDF Formula

- Molecular formula: total number of documents.
- Denominator: the number of documents containing the term.
- The result of the calculation is obtained by quoting the logarithm of the number of documents with a base of 10.

(3) Levenshtein

The Russian scientist Vladimir Levenshtein first proposed this algorithm in 1965 [24]. The basic form of Levenshtein is carried out using a regressive algorithm, where a threshold can be set as an upper limit for the number of steps to be moved. The Levenshtein distance formula is shown in Equation (4).

$$lev_{a,b}(i,j) = \begin{cases} \max(i,j) & \text{if } \min(i,j) = 0, \\ \min \begin{cases} lev_{a,b}(i-1,j) + 1 \\ lev_{a,b}(i,j-1) + 1 \\ lev_{a,b}(i-1,j-1) + 1_{(a_i \neq b_j)} \end{cases} & \text{otherwise.} \end{cases} \tag{4}$$

Equation (4) Levenshtein Distance Formula

(4) Jaccard

The intersection and union of the two samples can be used to derive the Jaccard similarity coefficient and Jaccard distance for different applications [25]. Jaccard’s coefficient gives the degree of similarity and the ratio between the size of the intersection of two sets and the size of the union in a finite set of samples. The Jaccard index formula is shown in Equation (5).

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \tag{5}$$

Equation (5) Jaccard Index Formula

(5) Synonyms

Synonyms is an open-source package for natural language tasks in Python and maintained by Chatopera. It provides a variety of NLP tasks, such as text alignment, recommendation algorithms, similarity calculation, semantic shifting, keyword extraction, concept extraction, automatic summarization, and search engines with a multisource lexical database for predata use. Regarding the word vector conversion task, the suite uses Google’s gensim suite with a word2vec model for conversion and the vector distance of words with a smooth gradient descent algorithm for approximation [26].

3.2.2. Step 2. Common Voting Mechanism

In this study, the same approximation task was assigned to the abovementioned five different algorithms, and after obtaining the best dish selection results for each algorithm, the highest vote was tallied as the best solution by pooling. The confidence scores of the algorithms were not equally comparable among the different algorithms (Levenshtein

distance does not have a confidence score, but a minimum step), as the meanings of the confidence scores of the algorithms are limited to intragroup comparisons. For this reason, instead of using the average of the sum of similar scores for the same project, the highest score of each algorithm was used for vote recognition, and in the final vote counting process, the votes for each algorithm were equal, which rendered it a fair majority vote decision.

3.2.3. Step 3. Nutritional Ingredient Analysis

The recipe data were obtained through a fuzzy analysis of the artificial intelligence model, and the nutritional ingredient analysis automatically determined all the ingredients in the dish. Finally, this study consolidated all the nutrients by means of portion calculation to complete the nutrient analysis. The dietary information conversion process is shown in Figure 2.

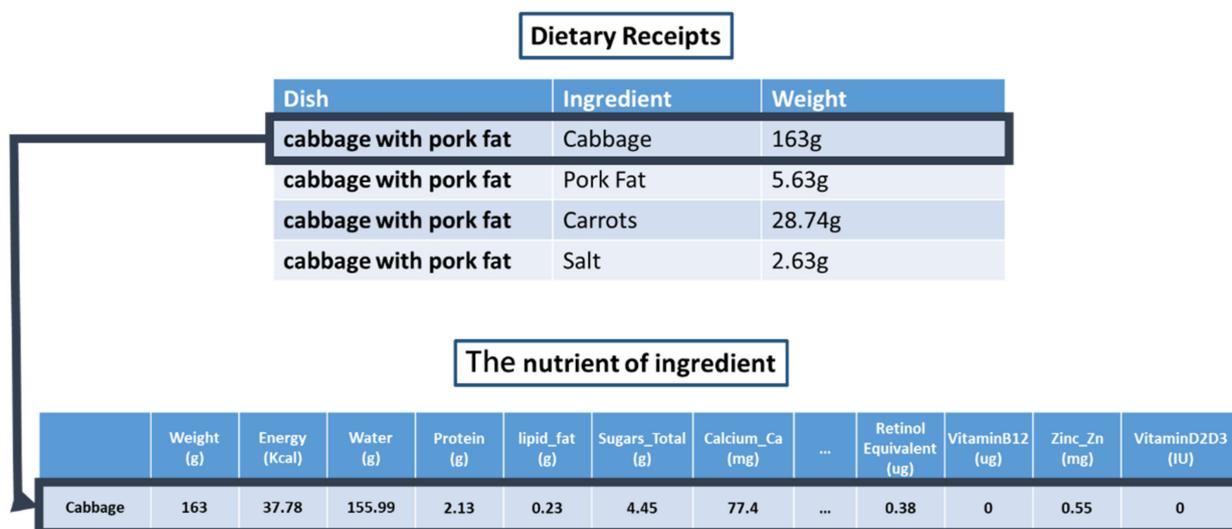


Figure 2. Nutrition Survey—Dietary Information Conversion Process.

4. Results

This study developed two models to form a precision nutrient analysis model. The first model is a Digitized Data Semantic Analysis Model for dish analysis and portion size determination. The second model is a Nutrient Analysis Model that uses five different algorithms to find precision recipes, which conducts analyses of dish ingredients and nutrients using a common voting process, and the final outputs from both models calculate the intake of 24 common nutrients. The operational framework of the model is illustrated below. The recipe database contains 1590 recipes and nutrient information for 7869 ingredients. The model operating framework is shown in Figure 3.

4.1. Operation Example

An example of a dietary recall record for precise nutritional analysis is as follows:

1. Input the dietary record to the model. Dietary Record: “Today I had a plate of cabbage with pork fat and a bowl of bamboo shoots and pork ribs soup.”
2. The names of the dishes and the portion sizes were analyzed by the Semantic Analysis Model. Nd is defined as time; Nf is defined as a quantity, and Na is defined as a common noun. Segmentation of record:

```
[{"label": "Today", "Pos": "Nd"}, {"label": "plate", "Pos": "Nf"}, {"label": "cabbage", "Pos": "Na"}, {"label": "pork fat", "Pos": "Na"}, {"label": "bowl", "Pos": "Nf"}, {"label": "bamboo shoots", "Pos": "Na"}, {"label": "pork ribs", "Pos": "Na"}, {"label": "soup", "Pos": "Na"}]
```

In this study, the plate of the dish represents 200 g, and the soup bowl represents 200 g.

3. Nutrition intake calculation by the Precision Nutrient Analysis Model.
 - (1) In Step 1: Each dish was separated into its ingredients according to the recipes. 200 g of cabbage with pork fat: Cabbage: 163 g, Pork Fat: 5.63 g, Carrots: 28.74 g, and Salt: 2.63 g. 200 g of bamboo shoots and pork ribs soup: Water: 50.13 g, White pepper: 0.63 g, Bamboo shoot: 73.43 g, Pork chops: 75.19 g, and Salt: 0.63 g.
 - (2) In Step 2: 24 nutrients were calculated for each ingredient, and the precision nutrient analysis results were calculated based on the sum of all nutrients.

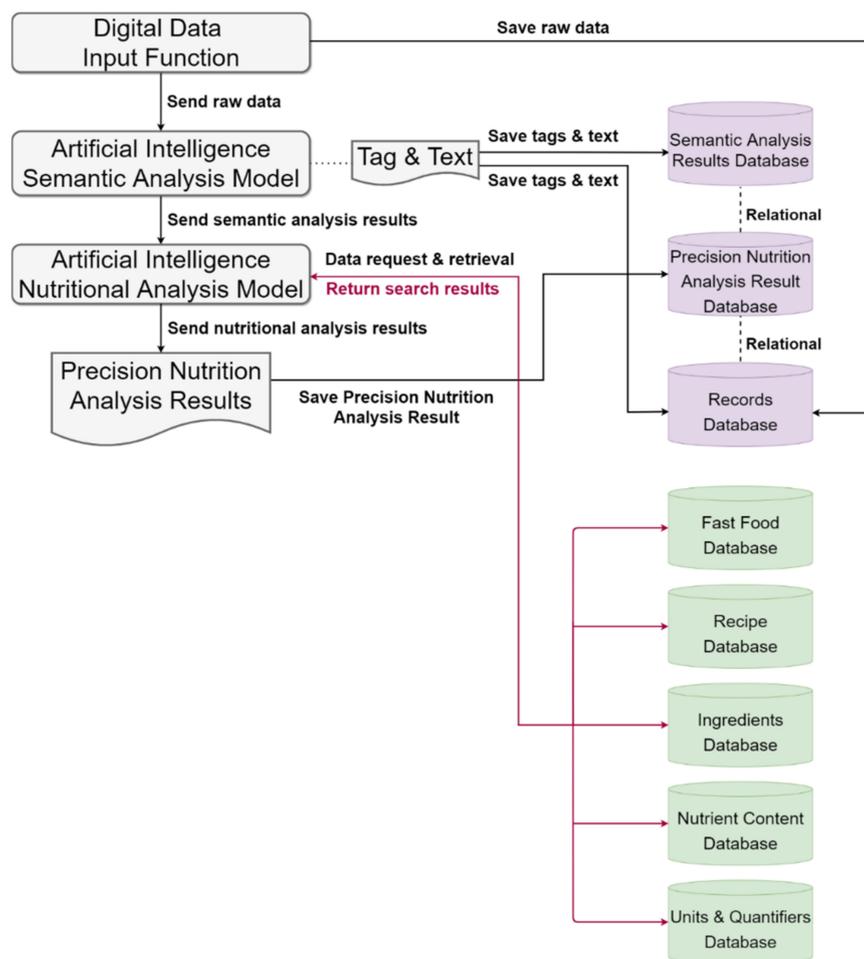


Figure 3. Model Operating Framework.

4.2. Model Accuracy Verification

The accuracy of the model was analyzed using data from the Nutrition Survey. In this study, the 2013–2016 National Nutrition Health Status Change Survey (NNHS) was used for analysis. The NNHS was initiated by the HPA MoHW and conducted in a four-year cycle and considered county and city distribution, as well as seasonal effects. The collected data were used as a reference for the formulation of national nutrition and health-related policies in Taiwan.

The aim of the survey is to understand the nutrition, health, diet, and lifestyle of the Taiwanese people and their relevance, in order to establish a long-term, stable, and nationally representative nutrition and health surveillance mechanism. The results can be used as a basis for government policies regarding diet and nutrition and health promotion and disease prevention and can help improve the health status of the population and prevent possible future health problems.

The NNHS uses a multistage stratified cluster sampling design, with the sample group being the entire age cohort, excluding pregnant and breastfeeding women, people without self-awareness, and institutional care residents, and the overall sample is representative of the Taiwanese population. The nutrition data were stored in a 24 h dietary memory record and analyzed by a professional nutritionist.

4.2.1. Data Resource

The “2013–2016 National Survey of Changes in Nutritional Health Status” was used to validate the accuracy of the model. The data contain a “24-h dietary recall nutrient intake sum analysis file” and a “24-h dietary recall food weight and nutrient ingredient file” with the information of 24 nutrients, including Energy, Water, Protein, Lipid Fat, Sugars Total, Calcium Ca, Phosphorus P, Iron Fe, VitaminB1, VitaminB2, Nicotinic, Vitamin C, Saturated Fat, Cholesterol, Vitamin E alpha TE, Sodium Na, VitaminB6, Magnesium, Dietary Fiber, Potassium K, Equivalent, VitaminB12, Zinc Zn, and VitaminD2D3.

- 24 h dietary recall nutrient intake sum analysis file (sum_nutrients_24hH)—total 2602 data entries. This file includes the data of the total nutrient intake in a single 24 h dietary recall survey.
- 24 h dietary recall food weight and nutrient ingredient file (food_wt_and_nutrients)—totaling 113,824 data entries. This file includes the data of the sum of nutrients for each individual dish, food, health product, etc.

4.2.2. Validation Process

- (1) Inputting data from the “102–105 National Survey of Changes in Nutritional Health Status” into a digitized data semantic analysis model;
- (2) Model analysis of dishes, portion sizes, and the ingredients in the dishes;
- (3) Analysis of nutrient intake using the AI Precision Nutrient Analysis Model;
- (4) Analyze the results against the “24-h dietary recall nutrient intake sum analysis file” and the “24-h dietary recall food weight and nutrient ingredient file”;
- (5) Compare the accuracy of the model.

4.3. Analysis Result

The results of the nutrition survey team analysis (from the 24 h dietary recall nutrient intake sum analysis file) were used as the gold standard, while the results of this study model analysis were used as the control group for the nutrient difference ratio analysis. The discrepancy comparison tables of the NNHS analysis with the results of this study in the 24 h dietary recall nutrient intake sum analysis are shown in Tables 1–3. A total of 2602 data entries were analyzed for total nutrient intake, with 24 different nutrients analyzed for each data item. The differences between the results of this study and the results of the nutrition survey are shown in Tables 1–3. While 13 nutrients had a total of more than 95% (2472 data entries) of the data with an intake error of <5%, 3 nutrients had a total of 90–94% of the data with an intake error of <5%; 5 nutrients had a total of 89.99–80% of the data with an intake error of <5%; Vitamin E alpha TE had a total of more than 95% (2472 data entries) with an intake error of <10%; Sugars Total and VitaminD2D3 had a 70% data error <10%.

The results of the nutrition survey team analysis (from 24 h dietary recall food weight and nutrient ingredient file) were used as the gold standard, while the results of this study model analysis were used as the control group for a nutrient difference ratio analysis. The discrepancy comparison table of the NNHS analysis with the results of this study in the 24 h dietary recall food weight and nutrient ingredients are shown in Tables 4–6.

A total of 113,824 data entries were analyzed for food weight and nutrient ingredients, with 24 different nutrients analyzed for each data item. The differences between the results of this study and the results of the nutrition survey are shown in Tables 4–6.

Table 1. Discrepancy comparison table of NNHS analysis with the results of this study in the sum_nutrients_24h (1).

| Error Range | Energy (kcal) | | Water (g) | | Protein (g) | | Lipid Fat (g) | | Sugars Total (g) | | Calcium Ca (mg) | | Phosphorus P (mg) | | Iron Fe (mg) | |
|-------------|---------------|-------|-----------|-------|-------------|-------|---------------|-------|------------------|-------|-----------------|-------|-------------------|-------|--------------|-------|
| <1% | 2187 | 84.1% | 2503 | 96.2% | 2284 | 87.8% | 1902 | 73.1% | 352 | 13.5% | 1168 | 44.9% | 2087 | 80.2% | 1401 | 53.8% |
| ≥1% and <2% | 297 | 11.4% | 68 | 2.6% | 228 | 8.8% | 319 | 12.3% | 436 | 16.8% | 601 | 23.1% | 306 | 11.8% | 734 | 28.2% |
| ≥2% and <3% | 45 | 1.7% | 24 | 0.9% | 43 | 1.7% | 136 | 5.2% | 270 | 10.4% | 298 | 11.5% | 82 | 3.2% | 199 | 7.6% |
| ≥3% and <4% | 22 | 0.8% | 5 | 0.2% | 12 | 0.5% | 84 | 3.2% | 186 | 7.1% | 194 | 7.5% | 44 | 1.7% | 116 | 4.5% |
| ≥4% and <5% | 9 | 0.3% | 1 | 0.0% | 10 | 0.4% | 42 | 1.6% | 142 | 5.5% | 109 | 4.2% | 21 | 0.8% | 39 | 1.5% |
| ≥5% and <6% | 8 | 0.3% | 1 | 0.0% | 8 | 0.3% | 29 | 1.1% | 91 | 3.5% | 68 | 2.6% | 15 | 0.6% | 41 | 1.6% |
| 6–10% | 27 | 1.0% | 0 | 0.0% | 12 | 0.5% | 56 | 2.2% | 316 | 12.1% | 122 | 4.7% | 32 | 1.2% | 51 | 2.0% |
| 11–15% | 6 | 0.2% | 0 | 0.0% | 2 | 0.1% | 13 | 0.5% | 172 | 6.6% | 22 | 0.8% | 8 | 0.3% | 7 | 0.3% |
| 16–20% | 1 | 0.0% | 0 | 0.0% | 2 | 0.1% | 6 | 0.2% | 136 | 5.2% | 11 | 0.4% | 3 | 0.1% | 7 | 0.3% |
| 21–30% | 0 | 0.0% | 0 | 0.0% | 0 | 0.0% | 8 | 0.3% | 170 | 6.5% | 2 | 0.1% | 2 | 0.1% | 7 | 0.3% |
| 31–40% | 0 | 0.0% | 0 | 0.0% | 1 | 0.0% | 6 | 0.2% | 91 | 3.5% | 5 | 0.2% | 2 | 0.1% | 0 | 0.0% |
| 41–60% | 0 | 0.0% | 0 | 0.0% | 0 | 0.0% | 1 | 0.0% | 121 | 4.7% | 0 | 0.0% | 0 | 0.0% | 0 | 0.0% |
| 61–80% | 0 | 0.0% | 0 | 0.0% | 0 | 0.0% | 0 | 0.0% | 49 | 1.9% | 2 | 0.1% | 0 | 0.0% | 0 | 0.0% |
| 81–100% | 0 | 0.0% | 0 | 0.0% | 0 | 0.0% | 0 | 0.0% | 32 | 1.2% | 0 | 0.0% | 0 | 0.0% | 0 | 0.0% |
| >100% | 0 | 0.0% | 0 | 0.0% | 0 | 0.0% | 0 | 0.0% | 0 | 0.0% | 0 | 0.0% | 0 | 0.0% | 0 | 0.0% |

Note: Each number indicates how much of the data are actually within that margin of error. The percentages indicate the percentage of the total data.

Table 2. Discrepancy comparison table of NNHS analysis with the results of this study in the sum_nutrients_24h (2).

| Error Range | VitaminB1 (mg) | VitaminB2 (mg) | Nicotinic (mg) | Vitamin C (mg) | Saturated Fat (g) | Cholesterol (mg) | Vitamin E Alpha TE (mg) | Sodium Na (mg) | | | | | | | | |
|-------------|-------------------|-------------------|-------------------|-------------------|----------------------|---------------------|----------------------------|-------------------|------|-------|------|-------|-----|-------|------|-------|
| <1% | 1990 | 76.5% | 1601 | 61.5% | 1811 | 69.6% | 1652 | 63.5% | 1440 | 55.3% | 1375 | 52.8% | 921 | 35.4% | 1836 | 70.6% |
| ≥1% and <2% | 345 | 13.3% | 644 | 24.8% | 416 | 16.0% | 234 | 9.0% | 480 | 18.4% | 586 | 22.5% | 398 | 15.3% | 296 | 11.4% |
| ≥2% and <3% | 100 | 3.8% | 135 | 5.2% | 146 | 5.6% | 139 | 5.3% | 193 | 7.4% | 542 | 20.8% | 307 | 11.8% | 106 | 4.1% |
| ≥3% and <4% | 59 | 2.3% | 81 | 3.1% | 76 | 2.9% | 81 | 3.1% | 112 | 4.3% | 20 | 0.8% | 240 | 9.2% | 87 | 3.3% |
| ≥4% and <5% | 35 | 1.3% | 38 | 1.5% | 45 | 1.7% | 60 | 2.3% | 76 | 2.9% | 12 | 0.5% | 176 | 6.8% | 60 | 2.3% |
| ≥5% and <6% | 17 | 0.7% | 17 | 0.7% | 23 | 0.9% | 37 | 1.4% | 69 | 2.7% | 8 | 0.3% | 128 | 4.9% | 55 | 2.1% |
| 6–10% | 40 | 1.5% | 46 | 1.8% | 57 | 2.2% | 104 | 4.0% | 137 | 5.3% | 21 | 0.8% | 288 | 11.1% | 93 | 3.6% |
| 11–15% | 7 | 0.3% | 23 | 0.9% | 20 | 0.8% | 50 | 1.9% | 49 | 1.9% | 12 | 0.5% | 82 | 3.2% | 37 | 1.4% |
| 16–20% | 4 | 0.2% | 6 | 0.2% | 5 | 0.2% | 35 | 1.3% | 11 | 0.4% | 9 | 0.3% | 31 | 1.2% | 12 | 0.5% |
| 21–30% | 4 | 0.2% | 9 | 0.3% | 1 | 0.0% | 52 | 2.0% | 21 | 0.8% | 3 | 0.1% | 18 | 0.7% | 8 | 0.3% |
| 31–40% | 1 | 0.0% | 2 | 0.1% | 0 | 0.0% | 27 | 1.0% | 11 | 0.4% | 1 | 0.0% | 9 | 0.3% | 8 | 0.3% |
| 41–60% | 0 | 0.0% | 0 | 0.0% | 1 | 0.0% | 40 | 1.5% | 2 | 0.1% | 1 | 0.0% | 4 | 0.2% | 1 | 0.0% |
| 61–80% | 0 | 0.0% | 0 | 0.0% | 1 | 0.0% | 27 | 1.0% | 1 | 0.0% | 1 | 0.0% | 0 | 0.0% | 1 | 0.0% |
| 81–100% | 0 | 0.0% | 0 | 0.0% | 0 | 0.0% | 11 | 0.4% | 0 | 0.0% | 5 | 0.2% | 0 | 0.0% | 2 | 0.1% |
| >100% | 0 | 0.0% | 0 | 0.0% | 0 | 0.0% | 53 | 2.0% | 0 | 0.0% | 6 | 0.2% | 0 | 0.0% | 0 | 0.0% |

Note: Each number indicates how much of the data are actually within that margin of error. The percentages indicate the percentage of the total data.

Table 3. Discrepancy comparison table of NNHS analysis with the results of this study in the sum_nutrients_24h (3).

| Error Range | VitaminB6 (mg) | | Magnesium (mg) | | Dietary Fiber (g) | | Potassium K (mg) | | Retinol Equivalent (ug) | | VitaminB12 (ug) | | Zinc Zn (mg) | | VitaminD2D3 (IU) | |
|-------------|----------------|-------|----------------|-------|-------------------|-------|------------------|-------|-------------------------|-------|-----------------|-------|--------------|-------|------------------|-------|
| <1% | 1511 | 58.1% | 1793 | 68.9% | 1518 | 58.3% | 1773 | 68.1% | 1234 | 47.4% | 1388 | 53.3% | 1579 | 60.7% | 723 | 68.9% |
| ≥1% and <2% | 518 | 19.9% | 529 | 20.3% | 327 | 12.6% | 487 | 18.7% | 475 | 18.3% | 350 | 13.5% | 626 | 24.1% | 209 | 20.3% |
| ≥2% and <3% | 180 | 6.9% | 153 | 5.9% | 196 | 7.5% | 155 | 6.0% | 241 | 9.3% | 255 | 9.8% | 206 | 7.9% | 162 | 5.9% |
| ≥3% and <4% | 123 | 4.7% | 34 | 1.3% | 95 | 3.7% | 76 | 2.9% | 148 | 5.7% | 165 | 6.3% | 75 | 2.9% | 143 | 1.3% |
| ≥4% and <5% | 60 | 2.3% | 22 | 0.8% | 61 | 2.3% | 43 | 1.7% | 98 | 3.8% | 95 | 3.7% | 27 | 1.0% | 120 | 0.8% |
| ≥5% and <6% | 48 | 1.8% | 18 | 0.7% | 53 | 2.0% | 20 | 0.8% | 61 | 2.3% | 68 | 2.6% | 18 | 0.7% | 86 | 0.7% |
| 6–10% | 95 | 3.7% | 22 | 0.8% | 187 | 7.2% | 35 | 1.3% | 140 | 5.4% | 181 | 7.0% | 35 | 1.3% | 339 | 0.8% |
| 11–15% | 28 | 1.1% | 14 | 0.5% | 60 | 2.3% | 8 | 0.3% | 98 | 3.8% | 53 | 2.0% | 18 | 0.7% | 250 | 0.5% |
| 16–20% | 11 | 0.4% | 5 | 0.2% | 36 | 1.4% | 5 | 0.2% | 39 | 1.5% | 12 | 0.5% | 8 | 0.3% | 75 | 0.2% |
| 21–30% | 16 | 0.6% | 3 | 0.1% | 46 | 1.8% | 0 | 0.0% | 48 | 1.8% | 13 | 0.5% | 0 | 0.0% | 81 | 0.1% |
| 31–40% | 8 | 0.3% | 4 | 0.2% | 10 | 0.4% | 0 | 0.0% | 9 | 0.3% | 5 | 0.2% | 1 | 0.0% | 54 | 0.2% |
| 41–60% | 3 | 0.1% | 5 | 0.2% | 9 | 0.3% | 0 | 0.0% | 5 | 0.2% | 4 | 0.2% | 1 | 0.0% | 58 | 0.2% |
| 61–80% | 0 | 0.0% | 0 | 0.0% | 3 | 0.1% | 0 | 0.0% | 4 | 0.2% | 2 | 0.1% | 4 | 0.2% | 40 | 0.0% |
| 81–100% | 1 | 0.0% | 0 | 0.0% | 1 | 0.0% | 0 | 0.0% | 0 | 0.0% | 5 | 0.2% | 0 | 0.0% | 31 | 0.0% |
| >100% | 0 | 0.0% | 0 | 0.0% | 0 | 0.0% | 0 | 0.0% | 2 | 0.1% | 6 | 0.2% | 4 | 0.2% | 231 | 0.0% |

Note: Each number indicates how much of the data are actually within that margin of error. The percentages indicate the percentage of the total data.

Table 4. Discrepancy comparison table of NNHS analysis with the results of this study in the food_wt_and_nutrients (1).

| Error Range | Energy (kcal) | Water (g) | Protein (g) | Lipid Fat (g) | Sugars Total (g) | Calcium Ca (mg) | Phosphorus P (mg) | Iron Fe (mg) | | | | | | | | |
|-------------|---------------|-----------|-------------|---------------|------------------|-----------------|-------------------|--------------|--------|-------|--------|-------|---------|-------|--------|-------|
| <1% | 97,435 | 85.6% | 99,390 | 87.3% | 97,211 | 85.4% | 102,695 | 90.2% | 95,184 | 83.6% | 97,059 | 85.3% | 100,940 | 88.7% | 99,540 | 87.5% |
| ≥1% and <2% | 6263 | 5.5% | 2592 | 2.3% | 4949 | 4.3% | 1878 | 1.6% | 3485 | 3.1% | 661 | 0.6% | 666 | 0.6% | 514 | 0.5% |
| ≥2% and <3% | 1177 | 1.0% | 1357 | 1.2% | 570 | 0.5% | 139 | 0.1% | 75 | 0.1% | 685 | 0.6% | 306 | 0.3% | 130 | 0.1% |
| ≥3% and <4% | 595 | 0.5% | 250 | 0.2% | 1114 | 1.0% | 67 | 0.1% | 265 | 0.2% | 146 | 0.1% | 101 | 0.1% | 88 | 0.1% |
| ≥4% and <5% | 153 | 0.1% | 67 | 0.1% | 514 | 0.5% | 338 | 0.3% | 370 | 0.3% | 915 | 0.8% | 663 | 0.6% | 1880 | 1.7% |
| ≥5% and <6% | 472 | 0.4% | 774 | 0.7% | 114 | 0.1% | 77 | 0.1% | 311 | 0.3% | 84 | 0.1% | 502 | 0.4% | 625 | 0.5% |
| 6–10% | 2231 | 2.0% | 1028 | 0.9% | 991 | 0.9% | 1096 | 1.0% | 1756 | 1.5% | 2583 | 2.3% | 4554 | 4.0% | 3749 | 3.3% |
| 11–15% | 575 | 0.5% | 2868 | 2.5% | 3500 | 3.1% | 850 | 0.7% | 258 | 0.2% | 878 | 0.8% | 399 | 0.4% | 791 | 0.7% |
| 16–20% | 790 | 0.7% | 606 | 0.5% | 122 | 0.1% | 518 | 0.5% | 285 | 0.3% | 4737 | 4.2% | 184 | 0.2% | 333 | 0.3% |
| 21–30% | 2952 | 2.6% | 1205 | 1.1% | 323 | 0.3% | 552 | 0.5% | 493 | 0.4% | 2020 | 1.8% | 666 | 0.6% | 547 | 0.5% |
| 31–40% | 383 | 0.3% | 1691 | 1.5% | 1596 | 1.4% | 2303 | 2.0% | 598 | 0.5% | 268 | 0.2% | 343 | 0.3% | 2657 | 2.3% |
| 41–60% | 80 | 0.1% | 1281 | 1.1% | 96 | 0.1% | 637 | 0.6% | 2803 | 2.5% | 760 | 0.7% | 504 | 0.4% | 1566 | 1.4% |
| 61–80% | 189 | 0.2% | 182 | 0.2% | 231 | 0.2% | 167 | 0.1% | 449 | 0.4% | 432 | 0.4% | 460 | 0.4% | 309 | 0.3% |
| 81–100% | 330 | 0.3% | 362 | 0.3% | 2142 | 1.9% | 1979 | 1.7% | 3087 | 2.7% | 2191 | 1.9% | 2051 | 1.8% | 378 | 0.3% |
| >100% | 199 | 0.2% | 171 | 0.2% | 351 | 0.3% | 528 | 0.5% | 4405 | 3.9% | 405 | 0.4% | 1485 | 1.3% | 717 | 0.6% |

Note: Each number indicates how much of the data are actually within that margin of error. The percentages indicate the percentage of the total data.

Table 5. Discrepancy comparison table of NNHS analysis with the results of this study in the food_wt_and_nutrients (2).

| Error Range | VitaminB1 (mg) | | VitaminB2 (mg) | | Nicotinic (mg) | | Vitamin C (mg) | | Saturated Fat (g) | | Cholesterol (mg) | | Vitamin E Alpha TE (mg) | | Sodium Na (mg) | |
|-------------|-------------------|-------|-------------------|-------|-------------------|-------|-------------------|-------|----------------------|-------|---------------------|-------|----------------------------|-------|-------------------|-------|
| <1% | 103,564 | 91.0% | 103,347 | 90.8% | 103,396 | 90.8% | 102,357 | 89.9% | 99,471 | 87.4% | 110,481 | 97.1% | 100,867 | 88.6% | 99,264 | 87.2% |
| ≥1% and <2% | 1899 | 1.7% | 1869 | 1.6% | 360 | 0.3% | 1721 | 1.5% | 782 | 0.7% | 27 | 0.0% | 150 | 0.1% | 3296 | 2.9% |
| ≥2% and <3% | 301 | 0.3% | 55 | 0.0% | 1075 | 0.9% | 237 | 0.2% | 280 | 0.2% | 1717 | 1.5% | 2102 | 1.8% | 227 | 0.2% |
| ≥3% and <4% | 376 | 0.3% | 388 | 0.3% | 197 | 0.2% | 69 | 0.1% | 1228 | 1.1% | 5 | 0.0% | 284 | 0.2% | 171 | 0.2% |
| ≥4% and <5% | 82 | 0.1% | 284 | 0.2% | 40 | 0.0% | 98 | 0.1% | 172 | 0.2% | | 0.0% | 31 | 0.0% | 138 | 0.1% |
| ≥5% and <6% | 98 | 0.1% | 257 | 0.2% | 175 | 0.2% | 62 | 0.1% | 151 | 0.1% | 10 | 0.0% | 119 | 0.1% | 240 | 0.2% |
| 6–10% | 2417 | 2.1% | 3394 | 3.0% | 2291 | 2.0% | 3267 | 2.9% | 1875 | 1.6% | 18 | 0.0% | 2710 | 2.4% | 4153 | 3.6% |
| 11–15% | 1331 | 1.2% | 578 | 0.5% | 609 | 0.5% | 308 | 0.3% | 782 | 0.7% | 17 | 0.0% | 129 | 0.1% | 503 | 0.4% |
| 16–20% | 216 | 0.2% | 691 | 0.6% | 877 | 0.8% | 337 | 0.3% | 1750 | 1.5% | 28 | 0.0% | 175 | 0.2% | 396 | 0.3% |
| 21–30% | 825 | 0.7% | 187 | 0.2% | 2583 | 2.3% | 621 | 0.5% | 259 | 0.2% | 29 | 0.0% | 2448 | 2.2% | 975 | 0.9% |
| 31–40% | 445 | 0.4% | 400 | 0.4% | 392 | 0.3% | 499 | 0.4% | 2312 | 2.0% | 40 | 0.0% | 203 | 0.2% | 1087 | 1.0% |
| 41–60% | 143 | 0.1% | 422 | 0.4% | 139 | 0.1% | 242 | 0.2% | 484 | 0.4% | 25 | 0.0% | 1142 | 1.0% | 639 | 0.6% |
| 61–80% | 186 | 0.2% | 95 | 0.1% | 77 | 0.1% | 392 | 0.3% | 411 | 0.4% | 27 | 0.0% | 335 | 0.3% | 231 | 0.2% |
| 81–100% | 1657 | 1.5% | 1478 | 1.3% | 1395 | 1.2% | 2022 | 1.8% | 2819 | 2.5% | 1201 | 1.1% | 2182 | 1.9% | 2305 | 2.0% |
| >100% | 284 | 0.2% | 379 | 0.3% | 218 | 0.2% | 1592 | 1.4% | 1048 | 0.9% | 199 | 0.2% | 947 | 0.8% | 199 | 0.2% |

Note: Each number indicates how much of the data are actually within that margin of error. The percentages indicate the percentage of the total data.

Table 6. Discrepancy comparison table of NNHS analysis with the results of this study in the food_wt_and_nutrients (3).

| Error Range | VitaminB6 (mg) | | Magnesium (mg) | | Dietary Fiber (g) | | Potassium K (mg) | | Retinol Equivalent (ug) | | VitaminB12 (ug) | | Zinc Zn (mg) | | VitaminD2D3 (IU) | |
|-------------|----------------|-------|----------------|-------|-------------------|-------|------------------|-------|-------------------------|-------|-----------------|-------|--------------|-------|------------------|-------|
| <1% | 99,965 | 87.8% | 93,126 | 81.8% | 105,576 | 92.8% | 102,151 | 89.7% | 101,920 | 89.5% | 109,024 | 95.8% | 97,681 | 85.8% | 110,264 | 96.9% |
| ≥1% and <2% | 294 | 0.3% | 2576 | 2.3% | 1991 | 1.7% | 452 | 0.4% | 121 | 0.1% | 9 | 0.0% | 151 | 0.1% | 346 | 0.3% |
| ≥2% and <3% | 2042 | 1.8% | 2091 | 1.8% | 295 | 0.3% | 734 | 0.6% | 91 | 0.1% | 1 | 0.0% | 964 | 0.8% | 3 | 0.0% |
| ≥3% and <4% | 295 | 0.3% | 160 | 0.1% | 458 | 0.4% | 162 | 0.1% | 59 | 0.1% | 4 | 0.0% | 1998 | 1.8% | 3 | 0.0% |
| ≥4% and <5% | 27 | 0.0% | 190 | 0.2% | 128 | 0.1% | 56 | 0.0% | 5 | 0.0% | 2 | 0.0% | 424 | 0.4% | 120 | 0.1% |
| ≥5% and <6% | 46 | 0.0% | 702 | 0.6% | 308 | 0.3% | 1118 | 1.0% | 24 | 0.0% | 12 | 0.0% | 95 | 0.1% | 86 | 0.1% |
| 6–10% | 2877 | 2.5% | 2045 | 1.8% | 1154 | 1.0% | 1391 | 1.2% | 3627 | 3.2% | 38 | 0.0% | 2568 | 2.3% | 14 | 0.0% |
| 11–15% | 1679 | 1.5% | 1406 | 1.2% | 296 | 0.3% | 131 | 0.1% | 118 | 0.1% | 1700 | 1.5% | 1084 | 1.0% | 1707 | 1.5% |
| 16–20% | 576 | 0.5% | 282 | 0.2% | 258 | 0.2% | 396 | 0.3% | 175 | 0.2% | 31 | 0.0% | 2991 | 2.6% | 10 | 0.0% |
| 21–30% | 1016 | 0.9% | 625 | 0.5% | 512 | 0.4% | 475 | 0.4% | 174 | 0.2% | 34 | 0.0% | 1764 | 1.5% | 29 | 0.0% |
| 31–40% | 205 | 0.2% | 300 | 0.3% | 63 | 0.1% | 239 | 0.2% | 1046 | 0.9% | 11 | 0.0% | 324 | 0.3% | 5 | 0.0% |
| 41–60% | 781 | 0.7% | 597 | 0.5% | 446 | 0.4% | 1962 | 1.7% | 407 | 0.4% | 171 | 0.2% | 643 | 0.6% | 12 | 0.0% |
| 61–80% | 574 | 0.5% | 1594 | 1.4% | 204 | 0.2% | 136 | 0.1% | 141 | 0.1% | 27 | 0.0% | 100 | 0.1% | 11 | 0.0% |
| 81–100% | 1896 | 1.7% | 7775 | 6.8% | 1878 | 1.6% | 3980 | 3.5% | 5097 | 4.5% | 2215 | 1.9% | 2532 | 2.2% | 1368 | 1.2% |
| >100% | 1551 | 1.4% | 355 | 0.3% | 257 | 0.2% | 441 | 0.4% | 819 | 0.7% | 545 | 0.5% | 505 | 0.4% | 52 | 0.0% |

Note: Each number indicates how much of the data are actually within that margin of error. The percentages indicate the percentage of the total data.

While 3 nutrients had a total of more than 95% of the data with an intake error of <2%, 9 nutrients had a total of 90–94% of the data with an intake error of <2%; 12 nutrients had a total of 89.99–80% of the data with an intake error of <2%.

5. Discussion

Each 24 h dietary recall nutrition survey in this study took approximately 40 min. The volume and complexity of the survey data and the variation in the ability to self-assess and recall portions can lead to individual subjective differences [13]. Similarly, the researchers or the methods used to collect dietary data may be biased [18].

Therefore, this study balanced the accuracy of nutrient intake analysis by compensating for errors through fuzzy analysis and artificial intelligence. Conventional FFQs are primarily designed to assess total nutrient intake or changes in intake over time [27–29]; however, the FFQ limits the range of foods that can be investigated as it combines food and beverages thus determining the exact amount of nutrients is less precise than other more detailed methods. It is also not possible to accurately measure absolute intakes of different food components. Moreover, FFQs require literacy and the physical ability to complete the questionnaire, and the FFQ survey can be burdensome for subjects and difficult or confusing to complete due to poor descriptions or difficult-to-understand questions. The most commonly used methods in nutrition research are the Diet Record, 24HR, and FFQ.

The Food Record is also used as the gold standard in validation studies [30]. Given the contingent nature of the respondents' food choices, a variety of food and beverage combinations [31] and nutrient supplementation [32] are the best methods to investigate. In order to reduce the burden on surveyors, the artificial intelligence model in this study has proven to be a feasible strategy for large-scale nutritional surveys after data discrepancy comparisons.

When comparing the difference between our model and the data analyzed in the actual nutrition survey, it was found that the results of the “24-h dietary recall food weight and nutrient ingredient” method were highly accurate, with less than 2% discrepancy in analysis for almost all nutrients. This result shows that the nutrients of the ingredient data in our model are correct. In the 24 h dietary recall nutrient intake sum analysis, the model was used to conduct an artificial intelligence analysis of the dishes, meaning it conducted an automated analysis of the components and servings to estimate nutrient intake. The results show a margin of error of less than 10% thus confirming the high accuracy of the model in this study.

6. Conclusions

This study proposed an Intelligence Precision Nutrient Analysis Model based on a digital data collection framework, where the nutrient intake was analyzed by entering dietary recall data. The AI Precision Nutrient Analysis Model was used to analyze the ingredients of the dishes and calculate nutrient intake by automatically analyzing the dishes, and portion sizes were analyzed using a digital data semantic analysis model. The results of this study show very little difference in nutrient intake between the model and the NNHS analysis and are highly accurate; therefore, the AI model can be used as a reference for nutrition surveys and personal nutrition analysis. In terms of data access, as there is not yet a complete set of publicly available data on food nutrient ingredients; more complete data and references on micro-nutrients should be available in the future. On the other hand, the scope of recipes should be expanded.

Author Contributions: The work presented in this paper was carried out in collaboration among all authors. H.-A.L. and C.-Y.L. formed the conception and study design. K.-W.C. and C.-Y.L. carried out the data analysis; H.-A.L. performed the literature review; C.-Y.L., T.-T.H., L.-H.Y., P.-H.W., and H.-H.K. performed the model development; H.-A.L., P.-H.W., and K.-W.C. drafted the manuscript, and C.-Y.H. made significant revisions and supplied valuable improvement suggestions. All authors have read and agreed to the published version of the manuscript.

Funding: This project received funding from the Ministry of Science and Technology, Taiwan, under the project no. 110-2221-E-227 -003 -MY3 and Ministry of Education, Taiwan, under the project no. 107EH12-402-022.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

References

- Cade, J.E. Measuring diet in the 21st century: Use of new technologies. *Proc. Nutr. Soc.* **2017**, *76*, 276–282. [CrossRef] [PubMed]
- Cade, J.E.; Consortium, O.B.O.T.D.; Warthon-Medina, M.; Albar, S.; Alwan, N.A.; Ness, A.; Roe, M.; Wark, P.A.; Greathead, K.; Burley, V.J.; et al. DIET@NET: Best Practice Guidelines for dietary assessment in health research. *BMC Med.* **2017**, *15*, 1–15. [CrossRef] [PubMed]
- Freedman, L.S.; Schatzkin, A.; Midthune, D.; Kipnis, V. Dealing with Dietary Measurement Error in Nutritional Cohort Studies. *JNCI* **2011**, *103*, 1086–1092. [CrossRef] [PubMed]
- Freedman, L.S.; Potischman, N.; Kipnis, V.; Midthune, U.; Schatzkin, A.; Thompson, F.; Troiano, R.P.; Prentice, R.; Patterson, R.; Carroll, R.; et al. A comparison of two dietary instruments for evaluating the fat–breast cancer relationship. *Int. J. Epidemiol.* **2006**, *35*, 1011–1021. [CrossRef]
- Carter, M.C.; Burley, V.; Nykjaer, C.; Cade, J. Adherence to a Smartphone Application for Weight Loss Compared to Website and Paper Diary: Pilot Randomized Controlled Trial. *J. Med. Internet Res.* **2013**, *15*, e32. [CrossRef]
- Timon, C.M.; Blain, R.J.; McNulty, B.; Kehoe, L.; Evans, K.; Walton, J.; Flynn, A.; Gibney, E.R. The Development, Validation, and User Evaluation of Foodbook24: A Web-Based Dietary Assessment Tool Developed for the Irish Adult Population. *J. Med. Internet Res.* **2017**, *19*, e158. [CrossRef]
- Boushey, C.J.; Harray, A.J.; Kerr, D.A.; Schap, T.; Paterson, S.; Aflague, T.; Ruiz, M.B.; Ahmad, Z.; Delp, E.J.; Harnack, L.; et al. How Willing Are Adolescents to Record Their Dietary Intake? The Mobile Food Record. *JMIR MHealth UHealth* **2015**, *3*, e47. [CrossRef]
- Thompson, F.E.; Dixit-Joshi, S.; Potischman, N.; Dodd, K.W.; Kirkpatrick, S.I.; Kushi, L.H.; Alexander, G.L.; Coleman, L.A.; Zimmerman, T.P.; Sundaram, M.E.; et al. Comparison of Interviewer-Administered and Automated Self-Administered 24-Hour Dietary Recalls in 3 Diverse Integrated Health Systems. *Am. J. Epidemiol.* **2015**, *181*, 970–978. [CrossRef]
- Thompson, F.E.; Byers, T. Dietary assessment resource manual. *J. Nutr.* **1994**, *124*, s2245–s2317. [CrossRef]
- Subar, A.F.; Kirkpatrick, S.I.; Mittl, B.; Zimmerman, T.P.; Thompson, F.E.; Bingley, C.; Willis, G.; Islam, N.G.; Baranowski, T.; McNutt, S.; et al. The automated self-administered 24-hour dietary recall (ASA24): A resource for researchers, clinicians, and educators from the national cancer institute. *J. Acad. Nutr. Diet.* **2012**, *112*, 1134–1137. [CrossRef]
- Campbell, V.A.; Dodds, M.L. Collecting Dietary Information from Groups of Older People. Limitations of the 24-Hr. Recall. *J. Am. Diet. Assoc.* **1967**, *51*, 29–33. [CrossRef] [PubMed]
- Heady, J.A. Diets of Bank Clerks Development of a Method of Classifying the Diets of Individuals for Use in Epidemiological Studies. *J. R. Stat. Soc. Ser. A* **1961**, *124*, 336. [CrossRef]
- Dwyer, J.T.; Gardner, J.; Halvorsen, K.; Krall, E.A.; Cohen, A.; Valadian, I. MEMORY OF FOOD INTAKE IN THE DISTANT PAST. *Am. J. Epidemiol.* **1989**, *130*, 1033–1046. [CrossRef] [PubMed]
- Iv, C.; Edgar, Godwin, S.L.; Vecchio, F.A. Cognitive strategies for reporting portion sizes using dietary recall procedures. *J. Am. Diet. Assoc.* **2000**, *100*, 891–897.
- Guthrie, H.A. Selection and quantification of typical food portions by young adults. *J. Am. Diet. Assoc.* **1984**, *84*, 1440–1444. [CrossRef]
- Bolland, J.E.; Yuhas, J.A.; Bolland, T.W. Estimation of food portion sizes: Effectiveness of training. *J. Am. Diet. Assoc.* **1988**, *88*, 817–821. [CrossRef]
- Howat, P.M.; Mohan, R.; Champagne, C.; Monlezun, C.; Wozniak, P.; Bray, G. Validity and reliability of reported dietary intake data. *J. Am. Diet. Assoc.* **1994**, *94*, 169–173. [CrossRef]
- Gibson, R.S. *Principles of Nutritional Assessment*; Oxford University Press: Oxford, UK, 2005.
- Kirkpatrick, S.I.; Dodd, K.W.; Tooze, J.; Bailey, R.L.; Freedman, L.; Midthune, D. *Measurement Error Webinar Series. Risk Factor Monitoring and Methods, National Cancer Institute, National Institutes of Health. 2012*; Bethesda: Bethesda, MD, USA, 2012.
- Gatford, S.E.R.; Walker, S.; Jones, S. Micheline Hancock-Beaulieu & Mike. In Proceedings of the Third Text RETrieval Conference (TREC 1994), Gaithersburg, MD, USA, 2–4 November 1994; Available online: https://trec.nist.gov/pubs/trec3/t3_proceedings.html (accessed on 1 January 2021).
- Jones, K.S.; Walker, S.; Robertson, S. A probabilistic model of information retrieval: Development and comparative experiments: Part 2. *Inf. Process. Manag.* **2000**, *36*, 809–840. [CrossRef]

22. Jones, K.S.; Walker, S.; Robertson, S. A probabilistic model of information retrieval: Development and comparative experiments: Part 1. *Inf. Process. Manag.* **2000**, *36*, 779–808. [[CrossRef](#)]
23. Jones, K.S. Index term weighting. *Inf. Storage Retr.* **1973**, *9*, 619–633. [[CrossRef](#)]
24. Levenshtein, V.I. Binary Codes Capable of Correcting Deletions, Insertions and Reversals. *Sov. Phys. Dokl.* **1966**, *10*, 707.
25. Michael, L.; Winter, D. Distance between Sets. *Nature* **1971**, *234*, 34–35.
26. Chatopera. Synonyms. Available online: <https://github.com/chatopera/Synonyms#references> (accessed on 20 January 2021).
27. Tucker, K.L.; Chen, H.; Hannan, M.T.; Cupples, L.A.; Wilson, P.W.F.; Felson, D.; Kiel, D. Bone mineral density and dietary patterns in older adults: The Framingham Osteoporosis Study. *Am. J. Clin. Nutr.* **2002**, *76*, 245–252. [[CrossRef](#)] [[PubMed](#)]
28. Wirfält, A.E.; Jeffery, R.W. Using Cluster Analysis to Examine Dietary Patterns: Nutrient Intakes, Gender, and Weight Status Differ Across Food Pattern Clusters. *J. Am. Diet. Assoc.* **1997**, *97*, 272–279. [[CrossRef](#)]
29. Haveman-Nies, A.; Tucker, K.; de Groot, L.; Wilson, P.; van Staveren, W. Evaluation of dietary quality in relationship to nutritional and lifestyle factors in elderly people of the US Framingham Heart Study and the European SENECA study. *Eur. J. Clin. Nutr.* **2001**, *55*, 870–880. [[CrossRef](#)]
30. Walter, W. *Nutritional Epidemiology*; Oxford University Press: Oxford, UK, 2012; Volume 40.
31. Freedman, L.S.; Midthune, D.; Arab, L.; Prentice, R.L.; Subar, A.F.; Willett, W.; Neuhouser, M.L.; Tinker, L.F.; Kipnis, V. Combining a Food Frequency Questionnaire With 24-Hour Recalls to Increase the Precision of Estimation of Usual Dietary Intakes—Evidence from the Validation Studies Pooling Project. *Am. J. Epidemiol.* **2018**, *187*, 2227–2232. [[CrossRef](#)]
32. Nicastro, H.L.; Bailey, R.; Dodd, K.W. Using 2 Assessment Methods May Better Describe Dietary Supplement Intakes in the United States. *J. Nutr.* **2015**, *145*, 1630–1634. [[CrossRef](#)]