Development and Validation of a Nutrient Profiling Model for Shopping Baskets: The Grocery Basket Score (GBS)

Sandro Marcon[1], André Naef[1], David Faeh[1,2,3], Paul Windisch[1*]

[1] dacadoo AG, Othmarstrasse 8, 8008 Zurich, Switzerland
[2] Epidemiology, Biostatistics and Prevention Institute, Division of Chronic Disease Epidemiology, University of Zurich, Zurich, Switzerland
[3] Health Department, Bern University of Applied Sciences, Bern, Switzerland

*** Correspondence:**
Paul Windisch, MD
dacadoo AG
Othmarstrasse 8
8008 Zurich
Tel: +41 44 251 23 23
ORCID-ID: 0000-0003-1040-4888
paul.windisch@dacadoo.com

Disclosures:
S.M., A.N., D.F., and P.W. report employment at or consulting for dacadoo AG, a technology company that has business activity with retailers, including grocery stores. dacadoo intends to file 'GBS' as a trademark and commercialize GBS™ software and/or consulting services. The other authors declare no additional conflict of interest.

Data availability:
This paper contains information from the Open Food Facts (https://world.openfoodfacts.org/) and OFF-GBS (https://github.com/windisch-paul/OFF-GBS) databases, which are made available under the Open Database License (ODbL, https://opendatacommons.org/licenses/odbl/1-0/).

Funding:
No external funding was received for this project.

Abstract

Background: Rating someone's diet as a whole as opposed to rating individual food items could help to better inform people about their nutrition and its health properties, e.g. by considering possible deficiencies. The goal of this project was to develop an automated rating of shopping baskets based on their health properties that can be deployed on top of existing loyalty or digital shopping basket programs by retailers without knowing for how long a basket will last or how many people consume it.
Methods: We developed a model that uses nutrient densities as opposed to absolute quantities based on self-reported food intake from the National Health and Nutrition Examination Survey (NHANES) as well as mortality follow-up data and conducted a validation against the Nutri-Score as well as the Alternate Healthy Eating Index (AHEI).
Results: The model we obtained punishes consuming calories from sugar and saturated fats, high salt intake, as well as consuming calories from beverages. Furthermore, a penalty is applied to a nutrition that is deficient in protein, vitamin C, or iron. Fiber consumption is rewarded. The model showed a high degree of correlation with the Nutri-Score (Pearson correlation coefficient: 0.60) and the AHEI (Pearson correlation coefficient: -0.62 for the AHEI without the protective effect of moderate alcohol consumption and -0.60 with it).
Conclusions: The proposed model is aligned with the recommendations formulated in nutritional guidelines, which is indicated by a high degree of correlation with the AHEI. Product-level profiling systems such as the Nutri-Score also exhibit a correlation with the model and can be used in combination.

Introduction

While what constitutes an optimal diet remains subject to discussion, the impact of diet on non-communicable diseases and mortality has been investigated in various studies.[1] As a consequence, nutrition profiling and rating systems have been developed to inform consumers and nudge them towards healthier shopping choices.[2]

One example is the Nutri-Score, which has been endorsed by several countries and is found on a variety of products.[3] Profiling systems such as the Nutri-Score focus on rating individual foods within a given group, which allows for printing scores on the products and enables the consumer to make direct comparisons between different items from the same category.[4] While this is certainly useful, focusing on individual foods also comes with tradeoffs. For example, one could compose an entire shopping basket consisting exclusively of items that received an A by the Nutri-Score due to their low calorie density and high fiber content while hardly containing any protein, even though the Nutri-Score rewards protein. As protein deficiency can be an issue in e.g., the elderly even in industrialized nations, it illustrates that assessing nutrition as a whole is more complex than summing up or even averaging individual food ratings. This is also relevant to consumers as they have to figure out what share of poorly rated products in their basket is still acceptable.

Nutritional guidelines such as the Alternate Healthy Eating Index (AHEI) tend to provide recommendations in quantities, such as servings, per day.[5,6] The AHEI was designed to predict chronic disease risk and has also been validated for predicting mortality.[7] While this approach can help consumers understand how they should generally compose their diet, it can still be challenging to translate this information into actionable shopping insights as a consumer would have to assess how many servings they are likely to get out of an item and keep track of the number of items they need for a given timeframe.

Therefore, the goal of this study was to develop a model that can be used to translate the comprehensive guidance from nutritional guidelines into actionable insights in the supermarket to complement product-level rating systems. To this end, the we set out to develop a model to rate shopping baskets for their health properties, assuming that these baskets account for a substantial amount of a consumer's total nutrition. While a single basket of food products may have little relevance, long term purchases from one retailer, or from a group of retailers who had accepted to share information, would be meaningful. Additionally, we did not want to develop a model that required a manual ingredient annotation but wanted a model that could be automatically deployed on top of existing loyalty or digital shopping basket programs by retailers. Finally, we sought a data-based model that avoided strong reliance on opinionated expert judgment and based the model on the link between grocery baskets and eventual mortality.

We conducted a validation by comparing the outputs of the final model with the Nutri-Score and the AHEI to ensure seamless integration into existing guidance.

The hypothesis was that while there would be differences to the Nutri-Score due to the factors mentioned above, both systems would be largely in agreement on the majority of items. For the AHEI, we hypothesized better agreement with our model than with the Nutri-Score due to the alignment of rating the consumer's food purchases as a whole (i.e. dietary patterns) as opposed to rating individual items.

Methods

Data

The data used in this study is a subset of the continuous National Health and Nutrition Examination Survey (NHANES)[8] from 2001 to 2018, which has been linked with the latest mortality data for the participants from 2019.[9] Therefore, the dietary data does not constitute data derived from retail purchases but is self-reported food intake data. As part of the NHANES nutritional assessment, participants of the respective cohort completed a 24-hour dietary recall interview in person with a trained dietary interviewer with another phone-based interview 3 to 10 days later. The food intake obtained from these recalls is linked through codes to the Food and Nutrient Database for Dietary Studies (FNDDS)[10], which was used to calculate the respective nutrient intakes for each participant in each recall.
Individuals younger than 17, older than 90, and those with missing nutritional information as well as those who reported no calorie intake were excluded from the analysis. This resulted in a dataset composed of 35'462 individuals, 25% of which were randomly assigned to the test set with the remaining individuals being assigned to the training set.

Considerations for Potential Dietary Signals

We constructed two types of dietary signals, namely "density signals" and "deficiency signals".
As one cannot know for how long a shopping basket will last in a household and how many people consume it, we decided to engineer features based on the density (i.e., the proportion) of a given nutrient relative to the total basket, which we will call "density signals". An example of such a density signal would be to calculate the calories from protein divided by the total calories in the basket, as opposed to absolute quantities like the amount of protein in grams.
For essential nutrients, such as proteins or certain vitamins and minerals, we also attempted their inclusion as "deficiency signals", i.e. by constructing a feature that indicates if a person is likely deficient for a given nutrient. To do so, we reviewed published nutrition guidelines and divided the daily recommended intake of the nutrient by the daily recommended calories from the same source, resulting in a recommended density of the nutrient per calorie. The deficiency signal is considered if the nutrient per calorie density in the basket is below the recommended density. The signal becomes stronger as the deficiency gets more severe.
All dietary signals were then evaluated as potential inputs to Cox proportional hazard models for predicting mortality.
As a requirement for all dietary signals, there should be a plausible biological mechanism linking each signal to mortality. For the protein example, such a link is that the body requires essential amino acids obtained from protein that it cannot synthesize itself and that a lack could, therefore, contribute to increased mortality. That link should be supported by a consistent correlation of that signal with mortality in NHANES[8].
Another consideration was that the signals should be computable for any retailer without additional manual labeling of foods being required. Therefore, we decided to focus on nutrients that are either part of the labeling requirements in many countries or that could be reliably imputed, e.g., from the name of the food and its categories. Lastly, we wanted the

model to be based on different, heterogeneous signals to reduce the likelihood of single items achieving very high scores by performing well across all signals, as a shopping basket consisting only of one item is unlikely to be conducive to healthy nutrition (low dietary diversity). In other words, we sought a mix of signals that provides a reasonable fingerprint of grocery shopping behavior.

Modeling

A schema of the modeling process is presented in Figure 1.
For the modeling and validation, we considered the items from the dietary recall as a person's shopping basket.
Due to high degrees of correlation between different potential signals, different combinations of signals while leaving out others were evaluated. The goal was to minimize redundancy, identify combinations of signals that fulfill all requirements mentioned above, avoid multicollinearity as a source of model misspecification, and ultimately to use the signals as inputs to a Cox proportional hazard model. Signals that were evaluated over the course of the project were calories from carbohydrates divided by total calories, calories from sugar divided by total calories, calories from proteins divided by total calories, calories from fat divided by total calories, calories from saturated fats divided by total calories, calories from unsaturated fats divided by total calories, calories from alcohol divided by total calories, amount of fiber divided by total calories, amount of sodium divided by total calories, calories from beverages divided by the weight of the beverages, protein deficiency, iron deficiency, and various vitamin deficiencies. In addition, we evaluated a joint model that used the same signals for both foods and beverages as well as separate models with different signals for foods and beverages respectively.
Standardization of the signals by subtracting the mean and dividing by the standard deviation was performed for interpretability reasons, such that the resulting model coefficients are on the same scale. Signals were also clipped to the limits of 1.5 times the interquartile range (IQR) rule for outlier detection (Tukey's fences).
Since the model's goal was to predict a time-to-event endpoint and the NHANES data is censored, Cox proportional hazard models were used to estimate the effect of different combinations of the nutrition signals defined above on mortality. The models were adjusted for age, sex, body-mass index (BMI) and the poverty-income ratio (i.e. household income divided by the poverty threshold) by including them as model coefficients and setting them to zero at inference time.
A Cox model that showed good performance at predicting mortality while considering a comprehensive set of signals was chosen for further validation and will be called the Grocery Basket Score (GBS) from here on out. The output that is obtained by multiplying the regression coefficient with the strength of the respective dietary signal is the logarithm of the hazard ratio. However, the number can, of course, be scaled to a more consumer friendly scale.
The separate levels of adjustment for the food and beverages model are available in Supplementary Tables 1 and 2, respectively.
Unless specified otherwise, data analysis was performed using Python (v. 3.8.10) as well as the numpy (v. 1.23.5), pandas (v. 1.3.5) and lifelines (v. 0.27.7) packages.

Validation

For the validation against the Nutri-Score, we retrieved food items from the Open Food Facts database.[11] In the preprocessing, we ignored items with a missing product code or name. Furthermore, we filtered items with missing energy information, missing Nutri-Score 2023, or where the macronutrients summed to more than 101 grams per 100 grams of food, items with saturated fat contents larger than total fat, items with more sugar than carbohydrates, items where one of the nutrients is larger than the 99th percentile and items missing information regarding any of protein, fat, carbohydrate, or salt content. The remaining 913'298 items were scored using the GBS and compared to the provided Nutri-Scores by computing the Pearson correlation coefficient.

Due to occasionally missing labeling information in the Open Food Facts database, we also developed an imputation model to complete the required label information. Specifically, our model is a deep neural network that implements a sequence model operating on the product categories and name using custom embeddings to impute fiber, iron, and vitamin C. A database with the imputed label information is available at https://github.com/windisch-paul/OFF-GBS.

In addition to computing the correlation coefficient for the whole dataset, clustering was performed to allow for a cluster-by-cluster correlation and comparison.

Items were clustered into different food clusters based on their assigned tags in the Open Food Facts database. Due to the high number of tags in the database and each item potentially being assigned multiple tags, it was deemed infeasible to manually define disjoint food clusters. A clustering approach based on context frequency of tags was used instead. As a first step, we filtered out tags that appeared less than 10'000 times in the whole corpus. This left 78 tags. As a second step, the term frequency-inverse document frequency matrix (tf-idf) was calculated. The idea was to vectorize tags based on their frequency and relevance. The tf-idf matrix was then reduced using singular-value decomposition (SVD) by keeping the 20 most relevant components. Finally, the K-Means clustering algorithm was applied to find 14 food clusters, with the 15th cluster manually assigned to beverages. The rationale of performing SVD is that K-Means would perform better on the reduced eigenvector space rather than on the original tf-idf matrix. The number of singular values and clusters was selected empirically using the "elbow-rule".

In addition to computing correlation coefficients, we calculated the absolute number and the percentage of items in a food cluster for which the difference between the z-scores of the GBS and the Nutri-Score was greater than 1, as well as the number and percentages of items with a difference greater than 2. This was done to highlight the items that caused the biggest disagreement.

For the validation against the AHEI, we computed the AHEI scores for 4'382 subjects with dietary recalls from NHANES 2017-18, computed the GBS using the same recalls and calculated the correlation coefficient between the GBS and AHEI, both with and without the protective effect of moderate alcohol consumption. The AHEI was computed in R (v. 4.3.2) using the dietaryindex package (v. 1.0.3). In addition, we also computed the Nutri-Score 2017 using the pyNutriScore package (v.0.0.2) and calculated a total Nutri-Score for each participant's nutritional information as a calorie-weighted average of the Nutri-Scores of the individual food items they consumed. We then calculated the correlation between the calorie-weighted Nutri-Scores and the AHEI.

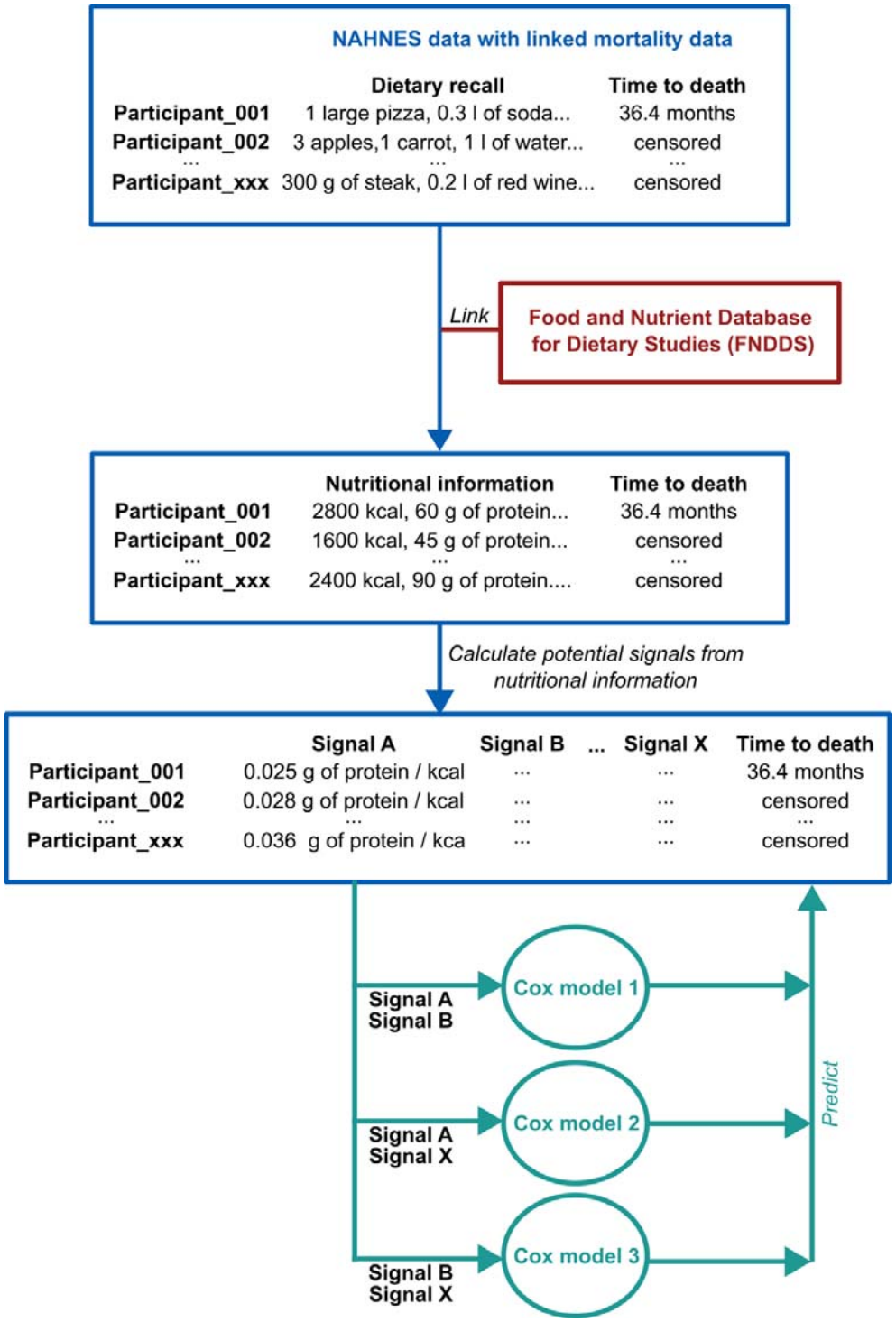

Figure 1. Schema of the modeling process. Dietary recall data from the National Health and Nutrition Examination Survey (NHANES) was linked with the Food and Nutrient Database for Dietary Studies (FNDDS) to obtain nutritional information for each participant. The dietary signals of different nutrients relative to the total dietary recall were used singly and in combinations as inputs to different Cox proportional hazard models to predict mortality. A model that showed good performance at predicting mortality while considering a comprehensive set of signals was selected for further validation.

Results

Modeling

After several iterations, a model that was obtained by scoring foods and beverages using separate models and combining the results as a calorie-weighted average of the respective scores demonstrated good performance at predicting mortality using a comprehensive set of signals with plausible regression coefficients. Thus, this model was selected as the first version of the Grocery Basket Score (GBS). Note that this selection represents a judgement call as selecting whichever combination of signals achieved the best performance could have yielded implausible coefficients due to multicollinearity.

A schema of doing inference with the GBS is provided in Figure 2. The signals and the regression coefficients are provided in Table 1. There are eight signals in total, one of which has a protective effect (amount of fiber divided by total calories), while the other signals are associated with increased risk. The biggest drivers of risk were vitamin C and iron deficiency as well as consuming a lot of calories via beverages, followed by a high intake of sodium, saturated fats, and sugar.

To derive additional insights into possible correlations between signals, we also created a correlation matrix between different parameters based on items in the Open Food Facts database, which is presented in Figure 3.

Validation

The overall correlation coefficient between the GBS and the Nutri-Score on single food items in the whole Open Food Facts database was 0.60. The clustering of items from the Open Food Facts database resulted in 15 clusters, and the correlation coefficients per cluster ranged from 0.25 to 0.75. The clusters and respective correlation coefficients are provided in Table 2.

Figure 4 shows the percentage of items per food cluster where the z-scores between the GBS and the Nutri-Score differ by more than 1 and 2 as well as the direction of the disagreement (i.e., which model rated more favorably). The lowest correlation coefficient was found in cluster 7 with a correlation coefficient of 0.25, which was due to items being scored more favorably by the Nutri-Score than by the GBS. In all other clusters, correlation coefficients were greater than 0.46. Samples of items scored differently by the GBS compared to the Nutri-Score are provided in Supplementary Tables 3 - 6.

The correlation coefficients between the GBS and the AHEI for 4’382 participants in NHANES 2017-2018 were -0.62 for the AHEI without the protective effect of moderate alcohol consumption and -0.60 with it. The correlation coefficients between the calorie-weighted Nutri-Scores and the AHEI were -0.48 and -0.47, respectively. The corresponding scatterplots are presented in Figure 5.

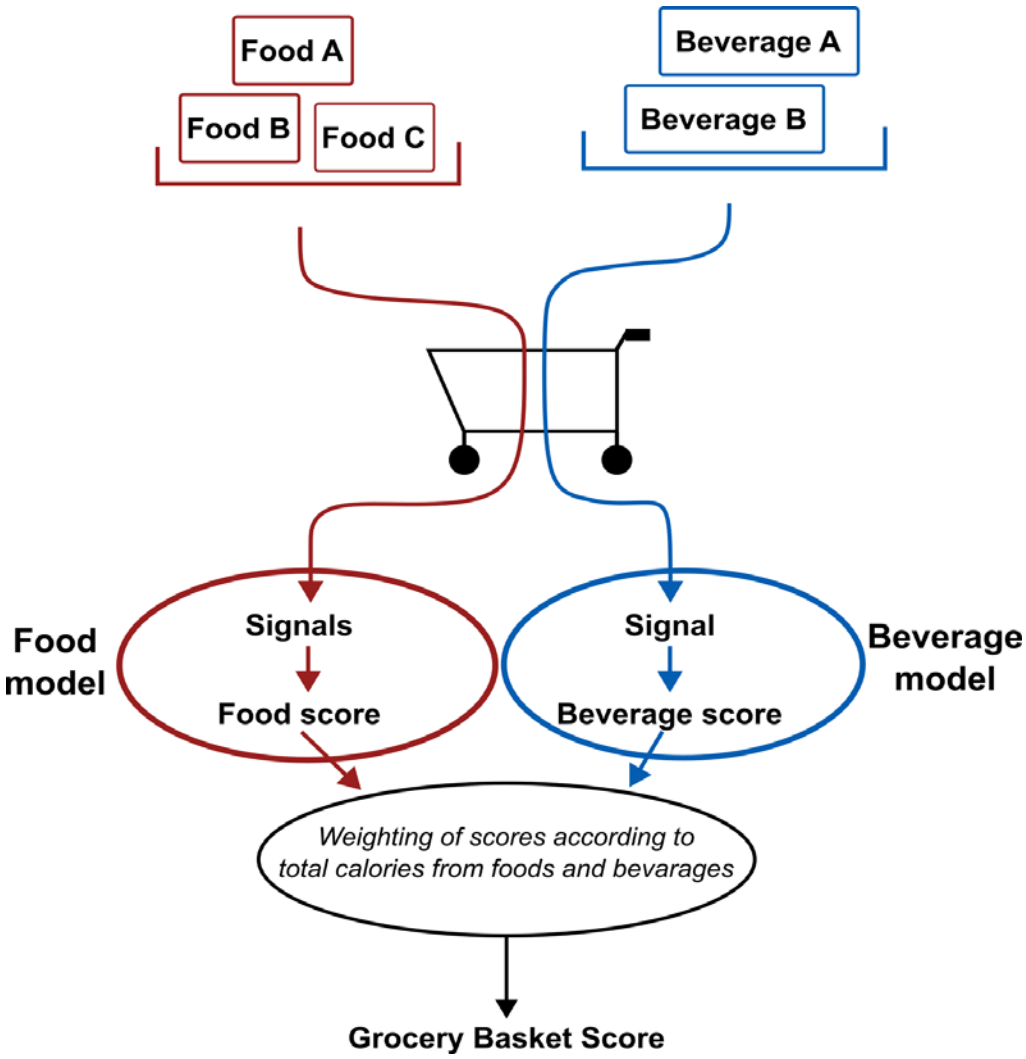

Figure 2. Schema of the Grocery Basket Score (GBS). The calories, nutrients, and weights in the basket are summarized for foods and beverages respectively. The data for beverages are passed to the beverages model which only computes one dietary signal (calories from beverages / weight of beverages). The data for the foods are passed to the food model which computes seven dietary signals (calories from saturated fats / total calories, protein deficiency, calories from sugar / total calories, amount of fiber / total calories, amount of sodium / total calories, vitamin C deficiency, iron deficiency).

The outputs of the food and beverage models are weighted by the ratio of the total calories. If a basket's calories come almost exclusively from food, the output of the food model will dictate the total score. If a basket's calories come almost exclusively from beverages, the output of the beverage model will dictate the total score.

Food model	
Dietary signal	**Regression coefficient**
Calories from saturated fats / total calories	0.059
Protein deficiency	0.020
Calories from sugar / total calories	0.056
Amount of fiber / total calories	- 0.081
Amount of sodium / total calories	0.064
Vitamin C deficiency	0.153
Iron deficiency	0.084
Beverage model	
Dietary Signal	**Regression coefficient**
Calories from beverages / weight of beverages	0.071

Table 1. Dietary signals and regression coefficients of the food and beverage models. Positive regression coefficients indicate increased mortality risk, while negative regression coefficients indicate decreased mortality risk. The deficiency signals were obtained by reviewing nutrition guidelines and dividing the daily recommended intake of the nutrient by the daily recommended calories from the same source, resulting in a recommended density of the nutrient per calorie. The deficiency signal is considered if the nutrient per calorie density in the basket is below the recommended density. The signal becomes stronger as the deficiency gets more severe. The output that is obtained by multiplying the regression coefficient with the strength of the respective dietary signal is the logarithm of the hazard ratio.

Cluster #	Open Food Facts Tags	Pearson correlation coefficient
0	seafood; fishes-and-their-products; fishes; fatty-fishes; canned-foods	0.62
1	fruits-and-vegetables-based-foods; plant-based-foods-and-beverages; plant-based-foods; vegetables-based-foods; canned-foods	0.66
2	biscuits-and-cakes; sweet-snacks; biscuits; snacks; cakes	0.74
3	meats-and-their-products; meats; prepared-meats; sausages; hams	0.73
4	dairies; fermented-foods; fermented-milk-products; cheeses; desserts	0.48
5	cereals-and-potatoes; plant-based-foods-and-beverages; plant-based-foods; cereals-and-their-products; breads	0.69
6	condiments; sauces; groceries; plant-based-foods-and-beverages	0.58
7	salty-snacks; appetizers; milks; chips-and-fries; crisps	0.25
8	meals; pasta-dishes; pizzas-pies-and-quiches; soups; meals-with-meat	0.46
9	frozen-foods; desserts; frozen-desserts; ice-creams-and-sorbets; dairy-desserts	0.58
10	confectioneries; sweet-snacks; snacks; cocoa-and-its-products; chocolate-candies	0.53
11	snacks; sweet-snacks; confectioneries; biscuits-and-cakes; biscuits	0.75
12	spreads; breakfasts; sweet-spreads; plant-based-spreads; fruit-and-vegetable-preserves	0.65
13	fats; vegetable-fats; vegetable-oils; plant-based-foods-and-beverages; plant-based-foods	0.72
14	beverages	0.60
Overall		**0.60**

Table 2. Pearson correlation coefficients by food clusters between the Grocery Basket Score (GBS) and the Nutri-Score. Greater absolute numbers indicate stronger positive or negative correlations, with 1 and -1 indicating the maximum positive and negative correlations and 0 indicating no correlation.

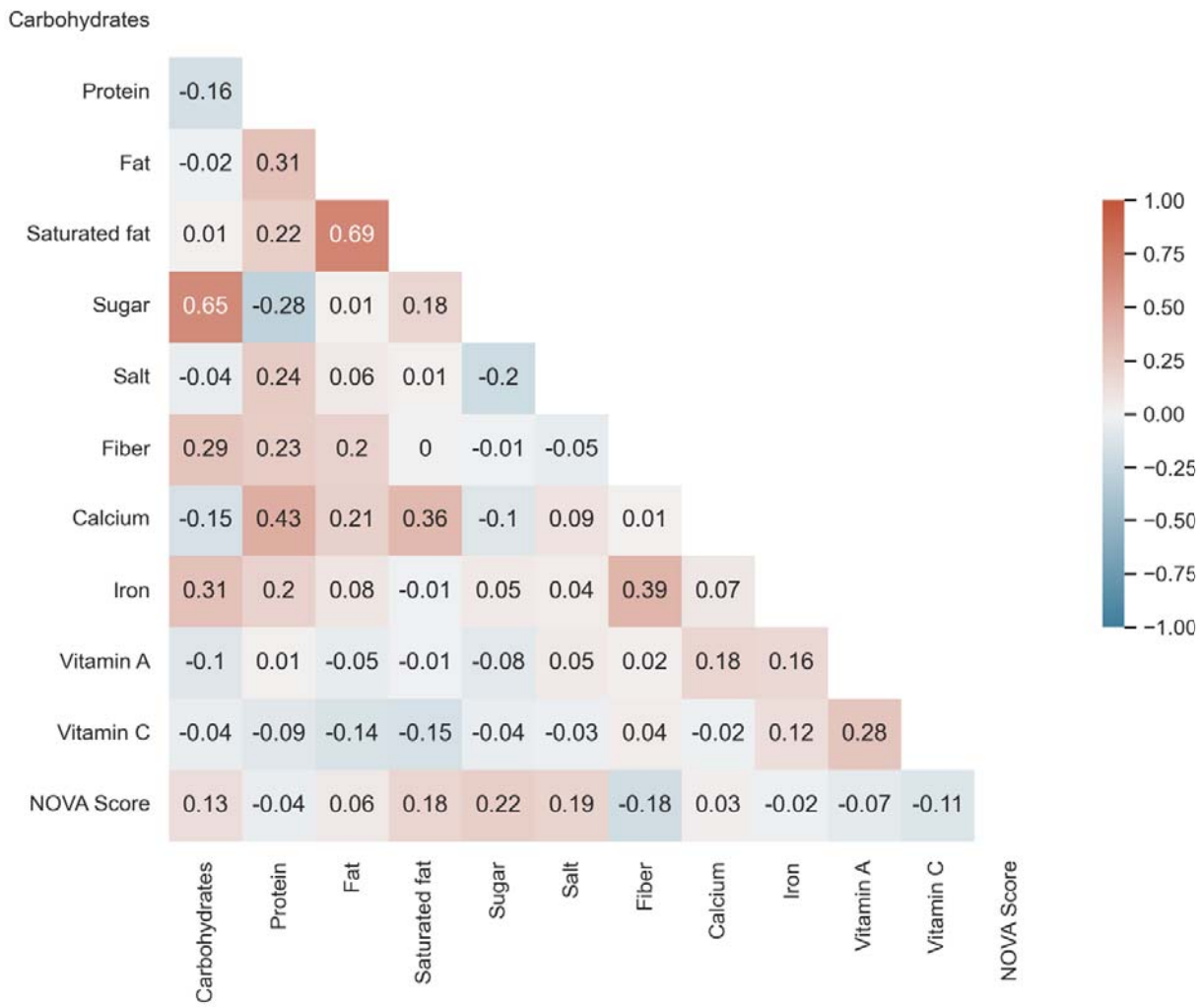

Figure 3. Pearson correlation coefficients between different parameters based on items in the Open Food Facts database.

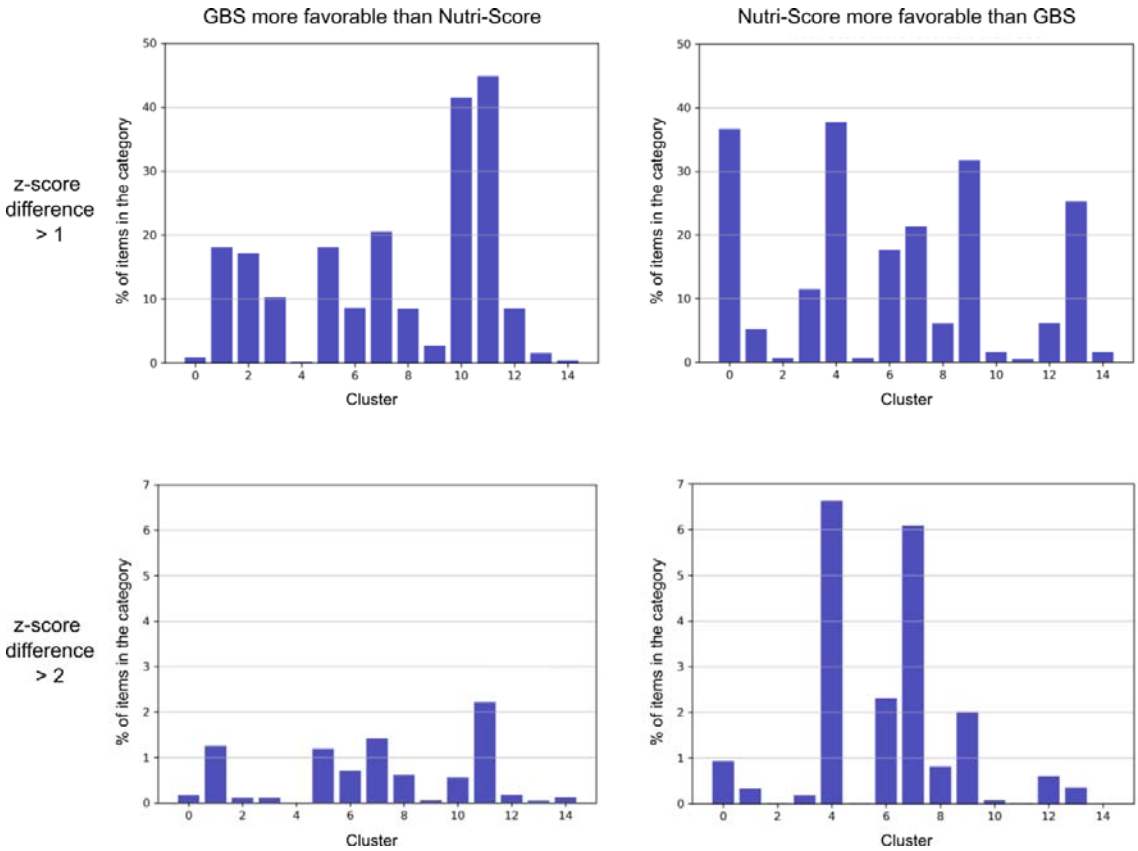

Figure 4. Outliers in Grocery Basket Score (GBS) and Nutri-Score ratings by food clusters. Percentage of items per food cluster where the z-scores between the GBS and the Nutri-Score differ by more than 1 (top) and 2 (bottom). Items that were rated more favorably by the GBS than the Nutri-Score are presented on the left, while items rated more favorably by Nutri-Score can be found on the right. The contents of the clusters can be found in Table 2.

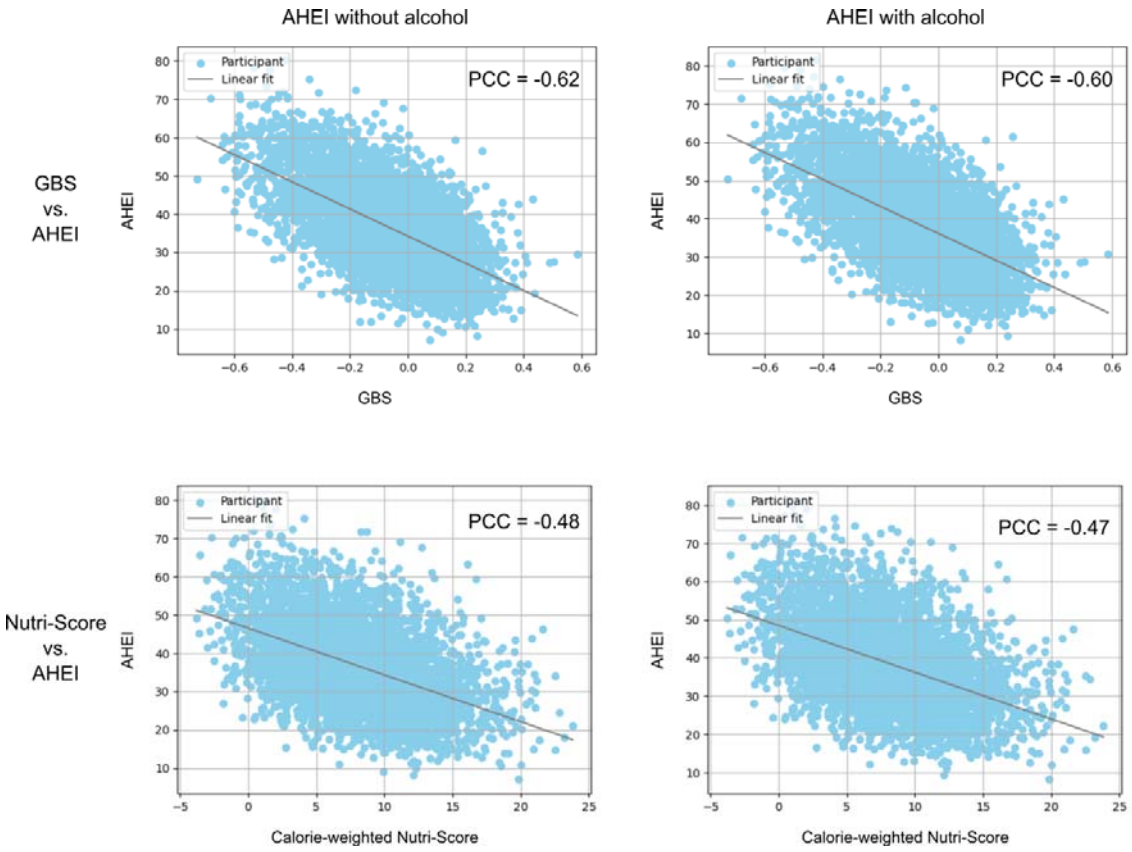

Figure 5. Scatterplot of the Grocery Basket Score (GBS - top) and calorie-weighted Nutri-Score (bottom) with Alternate Healthy Eating Index (AHEI) ratings. Each dot represents one of 4'382 NHANES 2017 - 2018 participants. Results of using the AHEI without the protective effect of moderate alcohol consumption are presented on the left and results with the protective effect of moderate alcohol consumption are displayed on the right. A steeper slope indicates a stronger correlation. *PCC = Pearson correlation coefficient.*

Discussion

We propose the Grocery Basket Score (GBS) as a method to automatically score foods and beverages in grocery shopping baskets according to their health properties based on their link with mortality. Its final score is computed as a calorie-weighted average of two separate models for foods and beverages, respectively.

The food model's characteristics can be derived from the regression coefficients assigned to its dietary signals: It punishes consuming calories from sugar and saturated fats, as well as high salt intake. Furthermore, a penalty is applied to baskets deficient in protein, vitamin C, or iron. Fiber consumption is rewarded. For the beverages model, calorie density is penalized. These regression coefficients of the GBS appear to be in alignment with popular nutrition guidelines issued, for example, by the U.S. Departments of Agriculture and of Health and Human Services or the World Health Organization.[12,13]

Even though the GBS was designed to be deployed on whole shopping baskets and not for assessing individual products, the high degree of correlation with the Nutri-Score, indicated by a correlation coefficient of 0.60 on items from the Open Food Facts database, demonstrates the considerable overlap in principles such as rewarding the presence of fiber and punishing the presence of saturated fats, sugar, and salt. This should reduce the risk of confusing consumers should the GBS be deployed in regions where the Nutri-Score is also used for individual products. The design differences to the Nutri-Score become evident when looking at the food clusters with lower correlation coefficients and the items for which the scores disagreed the most. The differences in the scoring of items in cluster 7 appear to be mainly driven by (low-fat) milk, which is rated more negatively by the GBS than by the Nutri-Score. The same applies to cluster 4 where the differences appear to be driven by dairy products such as cottage cheese or yoghurt. This is due to the lack of vitamin C, iron, and fiber, which is punished by our model as a balanced shopping basket should contain these nutrients. However, milk and egg whites are obviously not supposed to be sources of fiber. When part of a more comprehensive shopping basket, they would have a more positive effect due to reducing the protein deficiency and increasing the denominator (i.e., the total calories), thereby improving the sugar and saturated fat divided by total calorie signals. Cluster 10 contains sweet snacks such as chocolate which tend to be scored more positively by the GBS than by the Nutri-Score. This seems to be mainly because the Nutri-Score considers calorie density as a standalone signal and punishes items for having high calorie densities. The GBS takes a different route as it has been adjusted for BMI and does not consider calorie density as its own signal, effectively asking the question “If a consumer is able to maintain their weight, is it ok for them to consume this shopping basket”. While weight management is certainly an important goal when trying to compose a shopping basket, it is also relatively easy to track for the consumer when stepping on a scale. Therefore, the GBS focuses on issues that are more difficult to keep track of, such as the nutrient composition, which indirectly relates to weight management as fiber and protein tend to provide more satiety per calorie than sugars.

The high degree of correlation with the AHEI is promising and not surprising considering that the AHEI advises people to frequently eat fiber- and protein-rich foods while lowering sugar and salt consumption. It is also unsurprising that the correlation is even better when using the AHEI without assuming a protective effect of moderate alcohol consumption as the beverage model punishes drinks with higher calorie densities.

As a limitation, the fact that some signals are included in the final model while others are not, should not be interpreted as a claim that the included signals are somehow more predictive of mortality, let alone that there is more evidence for a causal relation, i.e., that supplementing the respective nutrients will improve mortality. It seems plausible that some of the signals used in the GBS function largely as proxies, such as the fiber signal and the vitamin C deficiency signal being proxies for fruit and vegetable or whole grains consumption which need to be considered when using the GBS to propose a healthy diet.

Furthermore, it would have been desirable to have even more signals to get a more comprehensive picture of someone's nutrition. However, due to the high degree of correlation between many signals, including more signals would not only provide, at best, minimal lift but also make the modeling more unstable due to variables stealing predictive power from each other.[14] While there are certainly other ways to do the modeling, and judgment calls of some degree were necessary to make the signal selection, we believe that the current model is capable of giving the consumer a useful evaluation of the quality of their diet, especially if grocery baskets scores from a single retailer, or groups of retailers, are combined over a period of time. The GBS is also relatively resistant to current attempts at gaming the system, e.g., by adding soluble fiber or proteins to beverages as this would not improve the beverage model that only focuses on calorie density. While having the degree of processing as a signal would have been interesting, this is harder to do without requiring manual labeling. In addition, ultra-processed foods are frequently high in salt, sugar, and saturated fats but relatively low in protein and fiber, as illustrated by the correlation matrix in Figure 3, so the degree of processing is at least indirectly represented in the GBS through surrogate signals.[15] This is supported by a recent study by Fang and colleagues, who found that adjusting for the AHEI attenuated a previously observed association of ultra-processed food intake with mortality.[16] The annotation problem and the representation through surrogates also apply to other potential signals like including whole grains, which have been suggested previously.[17]

Other limitations include the fact that this model was developed on a U.S. instead of a global population which might limit its ability to generalize to other parts of the world. In addition, the dietary information was self-reported. Furthermore, food items with missing nutritional information were not considered during the validation, which might have introduced bias, as well as the fact that for technical reasons Nutri-Score 2017 was used for the AHEI validation instead of the more recent version 2023. Lastly, another limitation of the GBS is that it does not model the negative health effects of alcoholic beverages that go beyond their calorie density.[18] The fact that adding the self-reported alcohol consumption as a signal did not increase the performance of the model when predicting mortality in NHANES might be due to inconsistencies in self-reported and actual alcohol consumption that have previously been described.[19]

A strength of the GBS is that it does not require manual annotation of food items prior to deployment, as it leverages the Open Food Facts database and an imputation engine. Since many retailers already have loyalty or digital shopping basket programs in place, this allows for automated real-world implementation. The use of high-level signals yielding a fingerprint as opposed to very granular lists of food items people should consume is aligned with the quality of evidence in the nutritional epidemiology space and allows for flexibility when composing a basket. Most importantly, the GBS gives users an automated assessment of their entire shopping basket without the need for manual tracking. Moreover, owing to its density-based nature, GBS does not require absolute quantities or knowledge of the household size. Even with shopping at multiple grocers and eating out, GBS can provide a

reasonable score as long as a representative subset of food intake is scored. As an outlook, the model could also be used to derive personal recommendations based on the strengths of the signals for a given basket.

In conclusion, the Grocery Basket Score (GBS) is a model for scoring shopping baskets that can consider possible deficiencies at the basket level in its ratings. The GBS scoring methodology is aligned with the recommendations formulated in nutritional guidelines, which is indicated by a high degree of correlation with the AHEI. Product-level profiling systems such as the Nutri-Score also exhibit a correlation with the GBS and can be used in combination.

References

1. Afshin A, Sur PJ, Fay KA, et al. Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2019;393(10184):1958-1972.

2. An R, Shi Y, Shen J, et al. Effect of front-of-package nutrition labeling on food purchases: a systematic review. *Public Health*. 2021;191:59-67.

3. Julia C, Etilé F, Hercberg S. Front-of-pack Nutri-Score labelling in France: an evidence-based policy. *Lancet Public Health*. 2018;3(4):e164.

4. van den Akker K, Bartelet D, Brouwer L, Luijpers S, Nap T, Havermans R. The impact of the nutri-score on food choice: A choice experiment in a Dutch supermarket. *Appetite*. 2022;168:105664.

5. McCullough ML, Willett WC. Evaluating adherence to recommended diets in adults: the Alternate Healthy Eating Index. *Public Health Nutr*. 2006;9(1A):152-157.

6. Chiuve SE, Fung TT, Rimm EB, et al. Alternative dietary indices both strongly predict risk of chronic disease. *J Nutr*. 2012;142(6):1009-1018.

7. Pestoni G, Karavasiloglou N, Braun J, et al. Does diet map with mortality? Ecological association of dietary patterns with chronic disease mortality and its spatial dependence in Switzerland. *Br J Nutr*. 2022;127(7):1037-1049.

8. Paulose-Ram R, Graber JE, Woodwell D, Ahluwalia N. The National Health and Nutrition Examination Survey (NHANES), 2021-2022: Adapting Data Collection in a COVID-19 Environment. *Am J Public Health*. 2021;111(12):2149-2156.

9. NCHS Data Linkage - Mortality Data - Public-Use Files. Published July 14, 2022. Accessed July 22, 2024. https://www.cdc.gov/nchs/data-linkage/mortality-public.htm

10. FNDDS : USDA ARS. Accessed July 22, 2024. https://www.ars.usda.gov/northeast-area/beltsville-md-bhnrc/beltsville-human-nutrition-research-center/food-surveys-research-group/docs/fndds/

11. Open Food Facts. Open Food Facts. Accessed April 28, 2024. https://world.openfoodfacts.org/discover

12. Snetselaar LG, de Jesus JM, DeSilva DM, Stoody EE. Dietary Guidelines for Americans, 2020-2025: Understanding the Scientific Process, Guidelines, and Key Recommendations. *Nutr Today*. 2021;56(6):287-295.

13. Healthy diet. Accessed April 29, 2024. https://www.who.int/news-room/fact-sheets/detail/healthy-diet

14. Kim JH. Multicollinearity and misleading statistical results. *Korean J Anesthesiol*. 2019;72(6):558-569.

15. Machado PP, Steele EM, Levy RB, et al. Ultra-processed food consumption and obesity in the Australian adult population. *Nutr Diabetes*. 2020;10(1):39.

16. Fang Z, Rossato SL, Hang D, et al. Association of ultra-processed food consumption with all cause and cause specific mortality: population based cohort study. *BMJ*. 2024;385:e078476.

17. Kissock KR, Vieux F, Mathias KC, et al. Aligning nutrient profiling with dietary guidelines: modifying the Nutri-Score algorithm to include whole grains. *Eur J Nutr*. 2022;61(1):541-553.

18. Wood AM, Kaptoge S, Butterworth AS, et al. Risk thresholds for alcohol consumption: combined analysis of individual-participant data for 599 912 current drinkers in 83 prospective studies. *Lancet*. 2018;391(10129):1513-1523.

19. Grüner Nielsen D, Andersen K, Søgaard Nielsen A, Juhl C, Mellentin A. Consistency between self-reported alcohol consumption and biological markers among patients with alcohol use disorder - A systematic review. *Neurosci Biobehav Rev*. 2021;124:370-385.