Author’s response to reviews

Title: Ability of the Nutri-Score front-of-pack nutrition label to discriminate the nutritional quality of foods in the German food market and consistency with nutritional recommendations

Authors:

Fabien Szabo de Edelenyi (f.szabo@eren.smbh.univ-paris13.fr)
Manon Egnell (m.egnell@eren.smbh.univ-paris13.fr)
Pilar Galan (p.galan@eren.smbh.univ-paris13.fr)
Nathalie Druesne-Pecollo (n.pecollo@eren.smbh.univ-paris13.fr)
Serge Hercberg (s.hercberg@eren.smbh.univ-paris13.fr)
Chantal Julia (c.julia@eren.smbh.univ-paris13.fr)

Version: 1 Date: 21 May 2019

Author’s response to reviews:

Reviewers’ comments followed by our point-by-point responses:

Reviewer #1:

(1) The manuscript is clearly written and well-structured. The paper contributes to the public health nutrition field by providing information on a topic currently being debated in Europe.

→ We are grateful for the Reviewer’s positive feedback on our manuscript.

I have few suggestions/queries to the authors:

(2) Abstract

A - Line 28: "categories" must be replaced by "food group".

→ The word has been corrected in the manuscript.

B - Line 29: How do the authors define "a high performance"? It would be good to report a definition in the Methods section of the abstract.
The ability of the front-of-pack nutrition label Nutri-Score in discriminating nutritional quality of foods and beverages was estimated by the number of available colours of Nutri-Score in each food group and sub-groups.

This has been stated in the method section of the manuscript, as follows:

“The ability of the FoPL to discriminate the nutritional quality of foods and beverages was estimated by the number of available colours of the Nutri-Score in each food group and sub-groups.”

(3) Methods

A - Food composition database subchapter:

The authors should specify whether the extracted data have been quality controlled.

Controls of extracted data were performed using manual check based on statistical control (outliers over P99) on individual variables used in the calculation of the Nutri-Score. Moreover, we also manually checked products with a mismatch between the energy calculated using carbohydrates, lipids and proteins contents and the energy variable in the database. These new checking were included during the revision process and explain the small differences in the results table.

All these elements were added to the manuscript, as follows:

→ “Controlled quality procedures included manual check based on outliers detection (over P99) on individual variables used in the calculation of the Nutri-Score in addition to controls already done at the OpenFoodFacts database level. Moreover, we also manually checked products with a mismatch between the energy calculated using carbohydrates, lipids and proteins contents and the energy variable in the database. Potential errors were corrected when possible using images available on OpenFoodFacts website. Otherwise the products were removed from analysis. Data were collected from 2012 to 2019, with regular updates each time a product is scanned again by a contributor.”

B - Moreover, was the composition of fruits/vegetables/pulses/nuts (%) included in the algorithm?

→ The composition of fruits / vegetables / pulses / nuts in percentage is included in the algorithm underlying the Nutri-Score (0 to 5 points are attributed in the score). When the data were available in the ingredients list of the food product, the percentage was then estimated. Otherwise, if one of the elements taken into account in the underlying algorithm was missing, the
Nutri-Score could not be calculated for the product. This has been added to the method section of the manuscript, as follows:

“Using crowdsourcing to collect food composition data of the food supply, specific data are collected by volunteer contributors including information about ingredients (including percentages of fruits and vegetable, legumes and nuts which are required for the computation of the Nutri-Score) and nutrition facts (including energy and mandatory nutrient-content per 100g: sugars, saturated fatty acids and sodium which are also used for the computation of the Nutri-Score) from foods products purchased in stores.”

C - The German FBDG, and corresponding reference, should be reported in the Methods section (maybe as a supplement to the manuscript). Moreover, are there any differences with the French guidelines?

→ The reference corresponding to the German food-based dietary guidelines has been added to the method section of the manuscript:


More details on the German FGBD were added in supplemental material.

We can notice that German FGBD are similar to the French or other European dietary guidelines. This has been added to the discussion section of the manuscript, as follows:

“In the French food environment, the classification of foods was overall consistent with French nutritional recommendations (which are very similar to German recommendations) and the discriminating ability of the 5 colours nutrition label was similar across food groups, within food groups and to a lower extent for equivalent foods from different brands.”

(4) Results

Figure 4: Food items should be sorted into food groups to improve clarity.

→ The Figure 4 has been modified according to the Reviewer 1’s comment.

(5) Discussion
A -Line 220: The consistency between the findings of the present study and those of the French studies should be made more explicit. Please provide data that allow to draw this conclusion.

→ The findings of the present study conducted in the German context were consistent with studies conducted in France. Indeed, as it was the case for Germany, in the French context, the classification of foods was consistent with French dietary guidelines and the discriminating ability of the 5 colours nutrition label (previous graphical format of the Nutri-Score) was similar across food groups, within food groups and to a lower extent for equivalent foods from different brands.

This was added to the discussion section of the manuscript, as follows:

“In the French food environment, the classification of foods was overall consistent with French nutritional recommendations (which are very similar to German recommendations) and the discriminating ability of the 5 colours nutrition label (previous graphical format of the Nutri-Score) was similar in France and Germany across food groups, within food groups and to a lower extent for equivalent foods from different brands.”

→ "Sub-group(s)" spelling should be made uniform throughout the document.

Reviewer #2:

(1) The introduction of front-of-pack labelling schemes is a hot topic in Europe. The performance of these labels (e.g. Nutri-Score) within country specific contexts is necessary to evaluate to choose the best option. The manuscript is of interest to contribute to the current debate on FOP labels in Germany but has some important limitations as well, in particular the use of crowd-sourced food database. The authors need to give more information in the methods on the collection period of the products included in the study, whether any checks are done to these data submitted by consumers and how representative these products are for all food products currently available on the German market. The latter could be done using Euromonitor data if no other data are available. In addition, some editing is necessary. There are numerous instances where different abbreviations for the same terms are used, and there is a lack of consistent word use.

→ We thank Reviewer 2 for his/her overall positive comment.
For the present study, the Open Food Facts database was used, coming from an international collaborative web project. Food composition data are constantly collected by volunteer contributors and made freely available as an open data source. For the present analyses, we used food composition data which were collected from 2012 to 2019 with regular updates each time a product is scanned again by a contributor. The quality of extracted data was verified using manual check based on statistical control (outliers over P99) on individual variables used in the calculation of the Nutri-Score. Moreover, we also manually checked products with a mismatch between the energy calculated using carbohydrates, lipids and proteins contents and the energy variable in the database. These new checking were included during the revision process and explain the small differences in the results table. However, quality controls are also performed at the OpenFoodFacts level.

Although this open data source allowed us to have access to a large sample of German foods and beverages currently available in supermarkets, we are conscious of the limitations of this database, and it has been stated in the limitations of the discussion section.

All these elements were added to the manuscript, as follows:

“Controlled quality procedures included manual check based on outliers detection (over P99) on individual variables used in the calculation of the Nutri-Score in addition to controls already done at the OpenFoodFacts database level. Moreover, we also manually checked products with a mismatch between the energy calculated using carbohydrates, lipids and proteins contents and the energy variable in the database. Potential errors were corrected when possible using images available on OpenFoodFacts website. Otherwise the products were removed from analysis. Data were collected from 2012 to 2019, with regular updates each time a product is scanned again by a contributor.”

[…]

“The main limitation of the study pertains to the use of the Open Food Facts database. Indeed, though the Open Food Facts database collects data from products currently available on the market directly from consumers, we were not able to analyze the representativeness of the sample of foods retrieved, either in terms of number of products or market share. However, our purpose was not to be exhaustive, but rather to test the performance of the Nutri-Score in real-life situations, for which the Open Food Facts database is sufficiently large to give a consistent evaluation.”

(2) Abstract:

A - Background
* Line 10: Be consistent in the way you describe 'foods/foodstuff/food products/…'. Please use the same word across the whole manuscript.

→ The terms have been harmonized throughout the manuscript.

B - Methods

* Line 19: Add years of data collection in the abstract

→ The years of data collection have been added in the abstract and the methods section of the manuscript as follows:

“Data were collected from 2012 to 2019, with regular updates each time a product is scanned again by a contributor.”

C - Results:

*The last line of the results section needs to describe the results underpinning this statement as the current sentence is an interpretation of results which are not described.

→ In the present study, it has been found that the nutrient profiling system underpinning the Nutri-Score was able to display the variability in nutritional quality of foods within the same food groups, with good discriminating performance (at least three colours represented with the Nutri-Score). The last sentence of the results section of the abstract was thus modified as follows:

“Moreover, we observed that the nutrient profiling system underpinning the Nutri-Score is able to display the variability in nutritional quality of foods within the same food groups, with good discriminating performance (at least three colours represented with the Nutri-Score).”

(3) Keywords:

*I would suggest to add 'food-based dietary guidelines' and 'front-of-pack labeling' to the keywords as these appear recurring in the manuscript

→ These keywords have been added to the manuscript accordingly.

(4) Background:
A - Please be consistent in the use of the abbreviations. FSA and NPS are used interchangeably, which is confusing for the reader.

→ These abbreviations have been modified for consistency throughout the manuscript.

B - The introduction could refer to a recent report by WHO Europe on FOP labelling in Europe:

→ We thank Reviewer 2 for this report; the reference has been added to the introduction section accordingly.

C - Line 55-57: Please clarify the range of healthiness of the color categories

→ The range of each category of the Nutri-Score has been added to Table 1 as a footnote, as follows:

“For foods: the FSAm-NPS score ranges from -15 to -1 points for the A category, from 0 to 2 for the B category, from 3 to 10 for the C category, from 11 to 18 for the D category, and 19 to 40 points for the E category.

For beverages: A corresponds to mineral waters exclusively. The FSAm-NPS score ranges from -15 to 1 point for the B category, from 2 to 5 for the C category, from 6 to 9 for the D category, and from 10 to 40 points for the E category.”

D - Line 82-86: Rewrite this sentence so it is clear to the reader that you are still talking about the 'international study' of the sentence before

→ The sentence has been modified as follows, to improve clarity:

“This international comparative experimental study aimed to compare the ability of five FoPLs [Nutri-Score, Australian Health Star Rating system (HSR), UK Multiple Traffic Lights (MTL), Chilean Warning labels and Reference Intakes (RIs) endorsed by manufacturers] to help consumers to understand the nutritional quality of different types of foods within different categories, in 12 countries including Germany.”

E - Line 82-86: Eliminate 'in 12 countries (including Germany)'

The sentence has been corrected accordingly.

F - Line 86-87: Replace 'the 12' by 'all'

The sentence has been corrected accordingly.

G - Line 86-87: Eliminate 'to'

The sentence has been corrected accordingly.

H - Line 92-94: Please rewrite so it is clear the Nutri-Score classification will be compared to 'German' FBDG

The sentence has been modified as follows:

“So, it appears of importance to assess how the Nutri-Score classifies foods in the German market and whether this classification aligns with the German food-based dietary guidelines (FBDG).”

(5) Methods: Food composition database

A - How many of the total number of products are current and is this an adequate reflection or representation of products currently on the German market? 8000 products seems a low number at any given time, and especially if the data are across several years. A check on representativeness could be done using Euromonitor data if no other data are available. Please give information on the years that these data were collected.

We agree that the number of foods present in the Open Food Facts database may appear low. Indeed, compared to the >250,000 foods present in the French database, the Germany database has not yet reached its full potential. However, our aim was to identify whether the Nutri-Score was able to discriminate the nutritional quality of common German foods and beverages, within and across food groups, rather than providing an exhaustive assessment of the nutritional quality of the German food supply.

The products that are referenced in the database were entered between 2012 and 2019, with a growing number of references added in 2017 and 2018. The references are updated by consumers when the products are modified in composition or when new packagings enter the market.
These elements were added to the manuscript, as follows:

“Data were collected from 2012 to 2019, with regular updates each time a product is scanned again by a contributor.”

B - Are there any checks on the data that are submitted by consumers?

Given the nature of the data collection, the quality of extracted data was verified using manual check based on statistical controls. Data submitted by consumers are also checked at the OpenFoodFacts level.

All these elements were added to the manuscript, as follows:

→ “Controlled quality procedures included manual check based on outliers detection (over P99) on individual variables used in the calculation of the Nutri-Score in addition to controls already done at the OpenFoodFacts database level. Moreover, we also manually checked products with a mismatch between the energy calculated using carbohydrates, lipids and proteins contents and the energy variable in the database. Potential errors were corrected when possible using images available on OpenFoodFacts website. Otherwise the products were removed from analysis.”

C - Did you recalculate the Nutriscore or just derived it from the OpenFoodFacts?

We did use in the paper the Nutriscore calculated by OpenFoodFacts. We calculated also the Nutri-Score using the data available in the OpenFoodFacts database to check whether the results were identical. We found a very high similarity between both calculated Nutri-Score (99.2%), with some differences coming from the estimate of fruits and vegetable contents. OpenFoodFacts uses powerful algorithms using ingredient list scan for example to estimate fruit and vegetables content of each product and it was then more effective to use directly the Nutri-Score available in the OpenFoodFacts database.

(6) Food classification

A - Be consistent in the way you describe 'food (sub)groups/food (sub)categories' as this can be confusing for the reader. Change also in the rest of the manuscript. Line 119-121: Rephrase 'missing group labelling'

→ The terms have been harmonized throughout the manuscript.

Lines 119-121 were also rephrased as follows:
“Foods for which the nutritional composition was incomplete for the computation of the Nutri-Score were also excluded (N=2781), as well as foods with missing food group (N=3289).”

B - Can you please explain the German food based dietary guidelines in more detail in the article

→ The reference corresponding to the German food-based dietary guidelines has been added to the method section of the manuscript:


More details on the German FGBD were added in supplemental material.

We can notice that German FGBD are similar to the French or other European dietary guidelines. This has been added to the discussion section of the manuscript, as follows:

“In the French food environment, the classification of foods was overall consistent with French nutritional recommendations (which are very similar to German recommendations) and the discriminating ability of the 5 colours nutrition label was similar across food groups, within food groups and to a lower extent for equivalent foods from different brands.”

C - Lines 119-121: Please specify how many products were excluded

→ 6070 products were excluded for incomplete information.

This was added to the methods section of the manuscript, as follows:

“Foods for which the nutritional composition was incomplete for the computation of the Nutri-Score were also excluded (N=2781), as well as foods with missing food group (N=3289).”

(7) Statistical analyses

A - Line 124-125: I do not agree with the use of 'FSAm/HCSP' for all different food categories, as this modified algorithm was made for cheeses, added fats and beverages only.

We agree that the modifications to the original algorithm were made only for three food groups, beverages, cheeses and added fats. However, the modifications also pertain to the thresholds used for the definition of the colours of the Nutri-Score, which differ from the original algorithm, for which only two thresholds have been defined. Therefore, consistently with previous
publications, we elected to rename the nutrient profile FSAm-NPS, to ascertain for readers that the modified algorithm corresponding to the underlying algorithm of the Nutri-Score was used.

B - Line 128-129: Please justify why the Nutri-Score performance was considered as 'good' when three or more colours were available in a food category. For some food sub groups it may be fine if this is not the case as some of these groups are considered unhealthy across the board (cfr the WHO Europe NP model)

→ The main goal of the Nutri-Score front-of-pack nutrition label is to show consumers the variability in nutritional quality between foods or beverages, and help them selecting healthier products between different alternatives in the same food group. This discriminating performance of the label therefore appears important whatever the food group. However, we agree that for some food categories that may be considered as ‘unhealthy’ across the board, this discriminating ability is of lesser importance. The Nutri-Score is however consistent with this approach, as we can notice in the present findings that less than 3 colours of Nutri-Score can be found in some specific sub-groups (e.g. sandwiches) which contain only D or E products, which is consistent with dietary recommendations. The chocolate group, for example, is almost entirely classified as E (93.3%), and the vast majority of biscuits and cakes, pastries or sweetened beverages are also classified as E.

(8) Results:

A - Line 138-142: The categorization of the method section is not corresponding with the one of the results section as the main category 'beverages' is missing.

→ The sentence has been modified as follows, in order to consider beverages:

“Concerning the German market, manufactured items with complete available data for the computation of the FSAm-NSP score in the Open Food Facts database were included in the analyses, corresponding to 8,587 foods and beverages: 527 products composed mainly of fruits and vegetables, 1,396 bread and cereal products, 688 meat, fish and eggs products, 1,875 milk and dairy products, 619 fats and sauces, 452 composite foods, 1,745 sugary snacks, 413 salty snacks, and 872 beverages.”

B - Line 183-186/187/189-190: The result section should be clear and concise, without any interpretation given of the results, as this belongs to the discussion section of the manuscript. Please rewrite.
The results have been rewritten as follows, in order to avoid any interpretation of the findings in this section of the manuscript:

“The distribution of the Nutri-Score within the different food groups and sub-groups is displayed in Table 1. A total of 79.7% of products from “fruits and vegetables”, 69.3% of products from “Cereals and potatoes” were classified as dark green (A) or green (B), while 93.4% of products from “Sugary snacks” were classified as orange (D) or dark orange (E). Among beverages, while a majority of fruit juices were classified as C (70.1%), soft drinks were classified as E.

Moreover, within almost each food group, differences in the nutritional quality of products between sub-groups were grasped by the Nutri-Score classification, with high discriminating ability (at least three colours represented as defined in the methods section). Thus, for example, within the “Milk and dairy products” sub-group, foods from the sub-group “Milk and yogurt” were mainly classified as products with higher nutritional quality – between dark green (A) and yellow (C) – than foods from “Ice creams” mainly categorized between yellow (C) and dark orange (D).

C - Line 186-187: Please clarify ‘a majority’ with percentages

→ A total of 70.1% of fruit juices were classified in the C category. The sentence has been modified accordingly:

“Among beverages, while a majority of fruit juices were classified as C (70.1%), soft drinks were classified as E.”

D - It may be nice to add some diagrams for some key food groups such as done in the Netherlands here: https://www.vmt.nl/Nieuws/Een_vijfde_tussendoortjes_scoort_goed_op_Nutriscore-190409072943?utm_source=dlvr.it&utm_medium=twitter

→ Pie charts for Sugary snacks, Pizza pies and quiches, Breakfast Cereals and Dairy desserts have been added in a new figure and a new sentence has been added in the text: “To illustrate the results from Table 1, pie charts for 4 key food groups (Breakfast cereals, Pizza pies and quiche, Dairy desserts and Sugary snacks) are shown in Figure 6.”

(9) Discussion:

A - Please provide a better comparison of the current results with the existing literature
We compared our results with previous French available data but, to our knowledge, there is no European results in the existing literature yet.

B - Could you also discuss the implications of the German situation for other countries currently having adopted Nutriscore (i.e. some supermarket chains are controlled in Germany like Aldi and Lidl)

The German situation regarding the implementation of the Nutri-Score in German supermarkets would have a direct impact on other countries. Indeed, from an economic and commercial point of view, the adoption of a single front-of-pack nutrition label in the different countries would be particularly important for industrialists and retailers exporting food products from and in Germany. This would be an important aspect, especially on the European food market.

Some elements were added in the conclusion of the manuscript, as follows:

“The German situation regarding the implementation of the Nutri-Score in German supermarkets would also have a direct impact on other countries, especially on the European food market. Indeed, the adoption of a single front-of-pack nutrition label in the different countries would be particularly important for industrialists and retailers exporting food products from and in Germany.”