

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

| | |
|----------------------------|---|
| TITLE (PROVISIONAL) | Perception of different formats of front-of-pack nutrition labels according to sociodemographic, lifestyle and dietary factors in a French population : cross-sectional study among the NutriNet-Santé cohort participants. |
| AUTHORS | Julia, Chantal; Péneau, Sandrine; Buscail, Camille; Gonzalez, Rebecca; Touvier, Mathilde; Hercberg, Serge; Kesse-Guyot, Emmanuelle |

VERSION 1 - REVIEW

| | |
|------------------------|---|
| REVIEWER | Jo Jewell WHO Regional Office for Europe |
| REVIEW RETURNED | 15-Feb-2017 |

| | |
|-------------------------|--|
| GENERAL COMMENTS | <p>This is a very relevant and clear paper, which will contribute to national and international policy discussions. The methodology focuses on a specific area relevant to use and utility of nutritional labels, notably consumer preference. So while the paper doesn't answer (nor does it attempt to answer) the bigger question of which labelling system is "best" - as this would require looking at many different aspects, including use and understanding - it does provide valuable additional perspectives on preference and perceptions of consumers in France. In addition to the range of preferences explore, the commitment to understanding different "clusters" of consumers is fully consistent with emerging evidence that consumers are varied and their dietary behaviours are influenced by different incentives and motivators. The disaggregation by age, sex, SES group and other variables such as adherence to dietary guidelines is relevant and interesting. The discussion is complete and well-referenced, thus contextualising the findings of this papers in the existing international literature. A particularly relevant point in the discussion relates to the importance of colour-coding for consumers and the demonstrated advantages of green-red as a cue, rather than other monochrome colouring that hold no association for consumers. I would recommend publication.</p> |
|-------------------------|--|

| | |
|------------------------|--|
| REVIEWER | Louise Signal University of Otago, Wellington |
| REVIEW RETURNED | 28-Feb-2017 |

| | |
|-------------------------|--|
| GENERAL COMMENTS | <p>Thank you for the opportunity to review this paper. The paper addresses an issue of considerable interest currently in many countries. It is appropriately researched and discussed. My only concern is that the writing should be simplified and paragraphs shortened.</p> |
|-------------------------|--|

| | |
|------------------------|---|
| REVIEWER | Erin Hobin Public Health Ontario, Canada |
| REVIEW RETURNED | 17-Mar-2017 |

| | |
|-------------------------|---|
| GENERAL COMMENTS | <p>General Comments:</p> <p>This article aimed to examine the perceptions of four front-of-pack nutrition labelling formats among adults in France, according to sociodemographic, lifestyle, and dietary factors. The authors conclude that the “Nutriscore” is most appealing to participants overall, specifically to participants with lower adherence to nutritional recommendations, and recommends this labelling system to policy-makers as the optimal choice. The paper is well written. The introduction provides an interesting overview of the food and nutrition strategy in the French context, and specifically the current debate in France related to various FOP labelling systems under consideration for nation-wide implementation. However, as outlined in detail below, a more thorough review of previous FOP studies and reviews is required in the introduction, more detail describing the study protocol and survey is required, the survey measures are relatively weak, the discussion was disjointed, and the conclusions seem somewhat overstated. Perhaps this paper could be revised and considered for a short communication.</p> <p>Introduction:</p> <ol style="list-style-type: none"> 1. This is a great first paragraph, highlighting the priority of preventing chronic disease industrialized countries and how is a key modifiable risk factor. Perhaps this paragraph (or first part of the first paragraph) can be expanded, first to emphasize the issues of chronic disease and poor diet internationally, then to include more French-specific data delineating the issues of chronic disease and diet-related conditions. Providing French-specific chronic disease and/or diet-related condition data in this first paragraph may then better set up discussing the FOP labelling formats in the French context in a second paragraph. 2. The introduction should contain a discussion of the review evidence on FOP labelling systems as well as the established strengths and weakness of current systems implemented in other jurisdictions. 3. Although no scientific study has directly compared the perceptions of the 4 specific FOP systems examined here, previous studies have compared various FOP systems containing similar design elements (e.g., multi-colour vs monochrome, single icon vs. ordinal scale). It would be worthwhile discussing these studies to help situate this work and what it contributes to the overall body of evidence. 4. Page 8, line 7: Why is the 5-CNL perceived as more effective than MTL or RIs? On what elements? In what context? In what tasks (e.g., identifying healthier foods, influencing actual purchases)? <p>Material and Methods</p> <ol style="list-style-type: none"> 5. Please use gender neutral pronouns throughout study (e.g., use “they” instead of “he” or “she”) 6. Do the participants complete the monthly surveys online? Are they required to complete a certain number of surveys per year to remain in the cohort? 7. Do the participants receive an incentive or any form of compensation for participating in the NutriNet-Sante cohort? 8. Was the FOP labelling questionnaire sent to all participants in the cohort in June 2016, the first questionnaire with survey items related to nutrition labelling, specifically FOP labels? How familiar were the |
|-------------------------|---|

| | |
|--|--|
| | <p>participants with the Nutriscore system versus the other systems before participating in this survey?</p> <p>9. Page 9, lines 18-19 – please consider revising sentence. For example, “A specific questionnaire was developed using survey items from previously published research to investigate participants’ perceptions of the four FOP labelling formats that are currently being debated within the French government (17;35).”</p> <p>10. Have these measures been validated? Were any other previously published survey measures used to test perceptions of FOP labelling systems considered? Outside of FOP labelling systems, many studies in the tobacco literature have used surveys to test participant perceptions of labels and maybe worthwhile to review.</p> <p>11. Does the survey assess level of nutrition knowledge? How was smoking status measured?</p> <p>12. Did all participants view the brief presentation of the four FOP labels before the start of the survey? Was the order in which the participants viewed or exposed to the four FOP labelling systems counter-balanced to avoid order effects?</p> <p>13. How were the four FOP labels displayed to participants? Were the labels displayed on products on screen (assuming the surveys are completed online)?</p> <p>14. Page 6, line 31: The acronym “EREN” is used before it has been defined in the paper.</p> <p>15. Please refer to the French FOP labelling system as either the 5 CNL or Nutriscore consistently throughout the paper. Additionally, it would be valuable to include more detail about how Nutriscore system works, and how it differs from other FOP labelling systems. What do the letters and colours represent?</p> <p>16. The three repeated 24-hour dietary records is definitely a strength of this study. Is this dietary data collected over a two-week period each year? Is this data collected online?</p> <p>17. Page 10, lines 50-51, please revise. For example, The dietary data from the 24h dietary record in the NutriNet-Sante study have been validated against interviewer-led dietary recalls conducted by trained dietitians, and against biomarkers of nutritional status.</p> <p>18. Please use a consistent name for the “questionnaire on the perceptions of FOP labels” throughout the paper. For example on Page 11, lines 4 to 12, various names are used to refer to the questionnaire in this one paragraph.</p> <p>19. Please either use the term sex or gender consistently throughout the paper.</p> <p>20. Please define acronyms first before using in the text (e.g., PNNS on page 11).</p> <p>21. Page 11, line 40. When the PNNS-GS is being described, is the term “moderation in consumption” referring to “portion sizes”?</p> <p>22. Page 11, line 43: why was adherence to physical activity recommendations not taken into account?</p> <p>23. Page 11, line 53: were energy expenditures based on physical activity data from the IPAQ used?</p> <p>Results</p> <p>24. Page 12, line 45: 38,604 participants completed the FOP questionnaire. What is the cooperation rate (e.g., of how many total participants)? Also, it would be valuable to include percentages (e.g., 21,702/38,604).</p> <p>25. A clear description of the positive and negative dimensions of perception is required. Moreover, how were the dimensions assessed? Is it a composite score?</p> <p>26. This is a minor suggestion, but I think it would be best to use a consistent number of decimal places throughout the paper.</p> |
|--|--|

| | |
|--|--|
| | <p>Discussion</p> <p>27. Page 15, lines 24 to 27: I am uncertain what this sentence is conveying. "Appealing, through complementary strategies, such as FOP nutrition labels, to those lacking the potential to translate nutritional recommendations or to those to which it would most benefit therefore appears of crucial importance."</p> <p>28. Page 16, lines 39 to 56: The finding related to colour, and readily interpretative colour-coding, is interesting.</p> <p>29. Based on the weaknesses of this study, including the cross-sectional study design, untested survey measures, and skewed sample, is it possible to conclude that the results of this study suggest the Nutriscore is a legitimate choice for policy-makers in France? I also question whether being more explicit that the authors of this study were involved in developing the Nutriscore system is warranted.</p> <p>Tables 2 and 3</p> <p>Please use consistent headers throughout. For example, does GDA refer to modified reference intakes?</p> |
|--|--|

VERSION 1 – AUTHOR RESPONSE

Reviewer 1: Jo Jewell (WHO Regional Office for Europe)

This is a very relevant and clear paper, which will contribute to national and international policy discussions. The methodology focuses on a specific area relevant to use and utility of nutritional labels, notably consumer preference. So while the paper doesn't answer (nor does it attempt to answer) the bigger question of which labelling system is "best" - as this would require looking at many different aspects, including use and understanding - it does provide valuable additional perspectives on preference and perceptions of consumers in France. In addition to the range of preferences explore, the commitment to understanding different "clusters" of consumers is fully consistent with emerging evidence that consumers are varied and their dietary behaviours are influenced by different incentives and motivators. The disaggregation by age, sex, SES group and other variables such as adherence to dietary guidelines is relevant and interesting. The discussion is complete and well-referenced, thus contextualising the findings of this papers in the existing international literature. A particularly relevant point in the discussion relates to the importance of colour-coding for consumers and the demonstrated advantages of green-red as a cue, rather than other monochrome colouring that hold no association for consumers. I would recommend publication.

We are very thankful for the very positive feedback the reviewer provided on the paper. We agree that our paper explores only one dimension of a front-of-pack label, and that complementary studies are necessary to ensure that other dimensions of the Nutriscore also ensure its efficiency (understanding and use in purchasing situations). However, a favourable perception is a pre-requisite for an efficient front-of-pack nutrition label, and as such, the Nutriscore appears as a legitimate option.

Reviewer 2: Louise Signal (University of Otago, Wellington)

Thank you for the opportunity to review this paper. The paper addresses an issue of considerable interest currently in many countries. It is appropriately researched and discussed. My only concern is that the writing should be simplified and paragraphs shortened.

We are thankful for this very positive feedback on our manuscript. During the revision process, we aimed at shortening paragraphs and simplifying the writing throughout the manuscript, as suggested.

Reviewer 3: Erin Hobin (Public Health Ontario, Canada)

General Comments:

This article aimed to examine the perceptions of four front-of-pack nutrition labelling formats among adults in France, according to sociodemographic, lifestyle, and dietary factors. The authors conclude that the "Nutriscore" is most appealing to participants overall, specifically to participants with lower

adherence to nutritional recommendations, and recommends this labelling system to policy-makers as the optimal choice. The paper is well written. The introduction provides an interesting overview of the food and nutrition strategy in the French context, and specifically the current debate in France related to various FOP labelling systems under consideration for nation-wide implementation. However, as outlined in detail below, a more thorough review of previous FOP studies and reviews is required in the introduction, more detail describing the study protocol and survey is required, the survey measures are relatively weak, the discussion was disjointed, and the conclusions seem somewhat overstated. Perhaps this paper could be revised and considered for a short communication.

We are thankful to the reviewer for the overall positive feedback on our manuscript and for the insightful comments on it. We have modified the manuscript according to the various points raised by the reviewer, and hope that the revised version offers a more synthetic and balanced overview of our work. In particular, we have strengthened the discussion and material sections, as suggested.

Introduction:

1. This is a great first paragraph, highlighting the priority of preventing chronic disease industrialized countries and how is a key modifiable risk factor. Perhaps this paragraph (or first part of the first paragraph) can be expanded, first to emphasize the issues of chronic disease and poor diet internationally, then to include more French-specific data delineating the issues of chronic disease and diet-related conditions. Providing French-specific chronic disease and/or diet-related condition data in this first paragraph may then better set up discussing the FOP labelling formats in the French context in a second paragraph.

As suggested, some data concerning the burden of chronic diseases in Western countries and France more particularly have been included in the first paragraph. It now reads:

Preventing non-communicable diseases has become a top priority for most industrialized countries, as they represent a major part of the burden of diseases (1). In France, cardiovascular diseases and cancers are the first causes of death, contributing each to approximately 30% of overall deaths every year (2). Diet has been recognized as a key modifiable factor which can influence – as preventive or risk factor – a wide range of non-communicable diseases, from cardiovascular disease to cancer, type 2 diabetes, metabolic syndrome or obesity (3-5).

2. The introduction should contain a discussion of the review evidence on FOP labelling systems as well as the established strengths and weakness of current systems implemented in other jurisdictions.

As suggested, the introduction was enriched in order to include background information on the implementation of FOP nutrition labels in the world and their typology. It now reads:

“Recently, novel complementary strategies have been put forward in a report to the French Minister of Health in 2014, highlighting the need for specific measures to modify the nutritional environment beyond the actions at the individual level (13). The report stressed in particular measures pertaining to nutrition labelling, in the form of a simplified front-of-pack (FOP) nutrition label, advertising regulation and nutritional taxation (13). Among the proposals of this report, the implementation of a FOP nutrition labelling system was considered as an effective opportunity by the Health Minister, and its principle was introduced in the 2016 French Health Law (14). Many countries have implemented FOP nutrition labels worldwide, either nutrient-specific, such as the ‘Multiple Traffic Light’ system in the UK (15) or summary measures, either simple – such as the Dutch ‘Choices’ logo (16) or the Nordic ‘Green Keyhole’ (17) – or graded – such as the ‘Health Star Rating System’ in New Zealand and Australia (18). Summary systems have been considered as more easily understood and interpreted than nutrient-specific labels, in particular for vulnerable populations (19;20). Moreover, colour-coded systems are considered more favourably perceived than monochrome systems (20;21).”

3. Although no scientific study has directly compared the perceptions of the 4 specific FOP systems examined here, previous studies have compared various FOP systems containing similar design elements (e.g., multi-colour vs monochrome, single icon vs. ordinal scale). It would be worthwhile discussing these studies to help situate this work and what it contributes to the overall body of evidence.

As suggested, the elements of comparison between summary and nutrient-specific labels, as well as colour-coded vs. monochrome have been added in the introduction section. Moreover, elements to

characterize the four labels compared in this study have been added. The introduction now reads: “Recently, novel complementary strategies have been put forward in a report to the French Minister of Health in 2014, highlighting the need for specific measures to modify the nutritional environment beyond the actions at the individual level (13). The report stressed in particular measures pertaining to nutrition labelling, in the form of a simplified front-of-pack (FOP) nutrition label, advertising regulation and nutritional taxation (13). Among the proposals of this report, the implementation of a FOP nutrition labelling system was considered as an effective opportunity by the Health Minister, and its principle was introduced in the 2016 French Health Law (14). Many countries have implemented FOP nutrition labels worldwide, either nutrient-specific, such as the ‘Multiple Traffic Light’ system in the UK (15) or summary measures, either simple – such as the Dutch ‘Choices’ logo (16) or the Nordic ‘Green Keyhole’ (17) – or graded – such as the ‘Health Star Rating System’ in New Zealand and Australia (18). Summary systems have been considered as more easily understood and interpreted than nutrient-specific labels, in particular for vulnerable populations (19;20). Moreover, colour-coded systems are considered more favourably perceived than monochrome systems (20;21). The initial report to the French Health Minister contained a detailed proposal for a simple colour-coded and graded label, supported by scientific studies (22-34) and independent government agencies evaluations (35;36) in the form of the 5-Colour Nutrition Label (5-CNL). However, alternative proposals were put forward during the debate by industry and retailers, in a vast lobbying campaign (37). Finally, four alternative formats emerged in the debate: the Nutriscore (an updated graphical version of the 5-CNL), SENS (a summary, graded and colour-coded label, developed and promoted by retailers), Multiple Traffic Lights (MTL, nutrient-specific and colour-coded label, currently used in the United Kingdom, UK) and a modified version of the Reference Intakes (mRIs, a nutrient-specific and monochrome label promoted by industry) (Figure 1).”

4. Page 8, line 7: Why is the 5-CNL perceived as more effective than MTL or RIs? On what elements? In what context? In what tasks (e.g., identifying healthier foods, influencing actual purchases)? As suggested, the ‘tasks for which the 5-CNL was considered more effective have been detailed. The sentence now reads:

“Some studies tend to indicate that the 5-CNL would be more favourably perceived than MTL or RIs (25) and that it may help consumers identifying (24;25) and purchasing healthier foods (26;34) but no data has been published on the mRIs or the SENS formats.”

Material and Methods

5. Please use gender neutral pronouns throughout study (e.g., use “they” instead of “he” or “she”) As suggested, pronouns have been modified throughout the manuscript

6. Do the participants complete the monthly surveys online? Are they required to complete a certain number of surveys per year to remain in the cohort?

All the surveys in the NutriNet-Santé study are carried online. Moreover, in order to limit as much as possible study dropouts and lost to follow-up, all surveys conducted after inclusion are considered as optional. Participants are provided with a measure of the importance of any given survey on a scale from 1 (low importance) to 3 (very important to research), in order to emphasize the relative importance of each survey for research purposes.

To clarify this point, the methods section was modified, as follows:

“Participants were selected from the NutriNet-Santé cohort. Briefly, the NutriNet-santé study is a prospective cohort study set in France in which inclusion and follow-up of volunteer participants are entirely performed on the Internet (38).”

[...]

“Once the subjects are included in the cohort, they receive monthly web-questionnaires pertaining to various aspects of dietary behaviour, physical activity and health, which are optional, and graded according to their relative importance for research.”

7. Do the participants receive an incentive or any form of compensation for participating in the NutriNet-Sante cohort?

Participants do not receive any incentives to participate in the online surveys.

This element was included in the Methods section, as follows:

“Participants do not receive any form of incentive or compensation to participate in the online surveys.”

8. Was the FOP labelling questionnaire sent to all participants in the cohort in June 2016, the first questionnaire with survey items related to nutrition labelling, specifically FOP labels? How familiar were the participants with the Nutriscore system versus the other systems before participating in this survey?

Indeed, all participants in the cohort in June 2016 received the questionnaire. This questionnaire was the second questionnaire pertaining to the perception of FOP nutrition labels, so participants were somewhat aware of the issue of FOP labelling. However, the formats that were presented in this specific study were somewhat different than those proposed in the previous survey. Moreover, more than a year passed between the two questionnaires.

For comparison purposes, please find below the four formats that were included in the previous survey and this study.

Figure 1 Formats presented in the first questionnaire on the perception of FOP nutrition labels

Figure 2 Formats presented in the second questionnaire on the perception of FOP nutrition labels

All formats were somewhat different between the first and the second questionnaire

- The 5-CNL presented in the first questionnaire is a previous version of the Nutriscore. The colour scale presented is not the same, as well as the global presentation
- The MTL presented in the first questionnaire does not report reference intakes, whereas the version presented in the second questionnaire corresponds to the updated 2013 version of the UK MTL
- The GDAs presented in the first questionnaires have been modified to include histograms of varying size according to the level of reference intakes
- The ‘Tick’ format was not included in the second questionnaire
- The SENS format was a novel format, proposed by retailers, that were not available at the time of the first questionnaire

Given all these elements, and the fact that no study was conducted solely on the Nutriscore, there appears to be no specific reason that the participants would be more sensitive to this FOP nutrition label. However, according to this comment, the discussion section was modified, as follows:

“Finally, the participants in the NutriNet-Santé study had already been involved in a previous survey on the perception of various FOP nutrition labels (25). However, the formats presented in the two versions of the questionnaire were somewhat different, and the delay between the two questionnaires of more than a year, therefore limiting the familiarity of the participants with the FOP nutrition labels formats displayed in this study. However, the participants were aware of FOP nutrition labelling, which could have affected their responses.”

9. Page 9, lines 18-19 – please consider revising sentence. For example, “A specific questionnaire was develop using survey items from previously published research to investigate participants’ perceptions of the four FOP labelling formats that are currently being debated within the French government (17;35).”

The sentence was modified as suggested:

“A specific questionnaire was develop using survey items from previously published research to investigate participants’ perceptions of the four FOP labelling formats that are currently being debated in France (25;42).”

10. Have these measures been validated? Were any other previously published survey measures used to test perceptions of FOP labelling systems considered? Outside of FOP labelling systems, many studies in the tobacco literature have used surveys to test participant perceptions of labels and maybe worthwhile to review.

The measures that were used to investigate the perception of FOP nutrition labels were previously used in several published papers, which took into account the literature on the subject of labelling perception. However, they were not validated stricto sensu. For future studies, a review of existing

questionnaires on perception, including the tobacco literature will be undertaken, in order to investigate the issue more thoroughly. However, it is not clear which elements could be considered as gold standard.

We agree that this may be a limitation of our survey, and have therefore modified the discussion as follows:

“Finally, the measures used in this study were not thoroughly validated but based on scientific literature. They derived from previously published work which took into account the literature on the perception of FOP nutrition labelling (42;61;73).”

11. Does the survey assess level of nutrition knowledge? How was smoking status measured?

The survey assessed nutrition knowledge using perceived self-reported knowledge by the participant, in four classes: ‘No’, ‘low’, ‘medium’ and ‘High’ nutritional knowledge. However, we elected not to include this measure as a covariate in our analyses, for several reasons: first, this measure was not strengthened by actual measures of nutrition knowledge and second, it appeared highly correlated with the other variables that were included in the model. However, according to this comment, a sensitivity analysis was conducted including this variable, which results are included below.

Nutriscore MTL SENS mRIs None P

43.23 27.31 17.05 7.31 5.10

Sex <0.0001

Men 47.58 44.25 43.16 42.07 42.12

Women 52.42 55.75 56.84 57.93 57.88

Age <0.0001

18-29 years-old 10.86 12.87 10.31 10.73 3.77

30-49 years-old 63.81 68.04 66.82 66.21 64.10

50-64 years-old 22.44 17.18 20.41 19.55 26.79

≥ 65 years old 2.89 1.91 2.46 3.51 5.34

Educational level <0.0001

Up to secondary 74.37 70.98 74.02 79.89 77.97

University, up to two years 14.18 15.02 14.50 11.42 11.93

University, ≥ 3 years 11.45 14.00 11.48 8.69 10.10

Income per consumption unit <0.0001

<1200/month 15.37 20.89 17.11 17.86 14.37

[1200 - 1800[€/month 34.22 33.31 37.19 37.49 37.42

[1800 - 2700[€/month 30.47 28.64 28.79 28.51 30.22

≥ 2700 €/month 19.94 17.16 16.91 16.14 17.99

Household composition <0.0001

Adults only 87.33 88.45 86.47 88.34 90.00

Adults and children 12.67 11.55 13.53 11.66 10.00

Smoking status <0.0001

Current smoker 10.32 10.68 9.41 8.81 15.17

Former smoker 30.09 33.39 32.44 34.04 31.72

Never smoker 59.59 55.93 58.15 57.15 53.11

Physical activity level <0.0001

High 32.10 34.10 31.42 35.81 30.10

Moderate 39.03 42.72 39.19 39.60 43.06

Low 28.87 23.18 29.39 24.59 26.84

mPNNs-GS <0.0001

Quartile 1 28.96 27.11 24.70 23.18 31.50

Quartile 2 25.28 21.70 26.05 19.77 29.81

Quartile 3 27.10 28.29 27.19 30.99 22.46

Quartile 4 18.66 22.90 22.06 26.06 16.23

Self-reported knowledge in nutrition <0.0001

High 10.97 15.25 12.66 16.78 13.34

Medium 51.10 56.69 48.88 51.53 35.50
Low 35.18 26.84 36.25 30.15 41.81
None 2.75 1.22 2.21 1.54 9.35

The inclusion of this variable did not modify the observed associations with the other covariates. Noteworthy, participants with the highest self-perceived nutrition knowledge were not more often in the Nutriscore cluster. We therefore elected not to include this analysis in the final manuscript. However, should the reviewer deem it necessary, we are willing to include this table as supplementary material.

Smoking status was measured using a self-administered online questionnaire, detailing the number of cigarettes, pipes and cigars consumed. Moreover, it detailed information on the date of quitting and consumption before quitting.

12. Did all participants view the brief presentation of the four FOP labels before the start of the survey? Was the order in which the participants viewed or exposed to the four FOP labelling systems counter-balanced to avoid order effects?

The questionnaire included several parts: 1) a presentation of the various formats, and the information they provided; 2) A task to classify various products (sets of three products) according to the FOP labels displayed (all participants tested all FOP labels, with a rotation on the order of the labels), in order to investigate objective understanding – this part was not used for the present analysis; 3) a questionnaire pertaining to the pragmatic legitimacy of labels (notions from marketing research), and showing each label – this part was not used for the present analysis; 5) the questionnaire on the perception of the various labels, in which the labels were again presented alone.

The order in which the labels were presented were the same for part 1, 3, 4 and 5. They were however displayed at random for part 2, which displayed the labels on the products.

According to this comment, the method section was modified, as follows:

“A specific questionnaire was developed using survey items from previously published research to investigate participants’ perceptions of the four FOP labelling formats that are currently being debated in France (25;42). The questionnaire also included other dimensions of FOP nutrition labelling evaluation (objective understanding and legitimacy), which were not used in this study.”

13. How were the four FOP labels displayed to participants? Were the labels displayed on products on screen (assuming the surveys are completed online)?

The labels were displayed in the questionnaire both on products and alone. In the specific part on perception, they were displayed alone.

14. Page 6, line 31: The acronym “EREN” is used before it has been defined in the paper.

As suggested, the acronyms have been now defined in their first occurrence. Moreover, we mentioned directly that the EREN team that developed the Nutriscore were the authors of the paper.

The paragraph now reads:

“Briefly, the Nutriscore, developed by the Nutritional Epidemiology Research Team (Equipe de Recherche en Epidémiologie Nutritionnelle, authors of this paper, EREN) scientific research team, and based on the British Food Standards Agency nutrient profiling system and adapted for the French context by the High Council for Public Health (1) presents for each food or beverage the overall nutritional quality on a 5-point colour-coded scale from Green to Red (Figure 1).”

15. Please refer to the French FOP labelling system as either the 5 CNL or Nutriscore consistently throughout the paper. Additionally, it would be valuable to include more detail about how Nutriscore system works, and how it differs from other FOP labelling systems. What do the letters and colours represent?

The 5-CNL refers to the previous version of the Nutriscore, which was somewhat different in terms of graphics and colours. Therefore, we maintained the distinction throughout the paper. We however made this clearer in the introduction, as follows:

“Finally, four alternative formats emerged in the debate: the Nutriscore (an updated graphical version of the 5-CNL), SENS (a summary, graded and colour-coded label, developed and promoted by retailers), Multiple Traffic Lights (MTL, nutrient-specific and colour-coded label, currently used in the

United Kingdom, UK) and a modified version of the Reference Intakes (mRIs, a nutrient-specific and monochrome label promoted by industry) (Figure 1).”

16. The three repeated 24-hour dietary records is definitely a strength of this study. Is this dietary data collected over a two-week period each year? Is this data collected online?

The 24h dietary records are based on three days randomly selected within a two-weeks period, including two week days and a week-end day. All data is collected online, and the online versions of the 24 records have been validated against interview by dietitians and biological markers of the nutritional status.

17. Page 10, lines 50-51, please revise. For example, The dietary data from the 24h dietary record in the NutriNet-Sante study have been validated against interviewer-led dietary recalls conducted by trained dietitians, and against biomarkers of nutritional status.

The sentence was modified as suggested. It now reads:

“The dietary data from the 24h dietary record in the NutriNet-Sante study have been validated against interviewer-led dietary recalls conducted by trained dietitians, and against biomarkers of nutritional status (33;41;42).”

18. Please use a consistent name for the “questionnaire on the perceptions of FOP labels” throughout the paper. For example on Page 11, lines 4 to 12, various names are used to refer to the questionnaire in this one paragraph.

As suggested, the questionnaire was termed consistently as ‘questionnaire on the perceptions of FOP labels’.

19. Please either use the term sex or gender consistently throughout the paper.

As suggested, sex was used preferentially over gender throughout the manuscript.

20. Please define acronyms first before using in the text (e.g., PNNS on page 11).

As suggested, the acronyms have been fully defined at their first appearance (PNNS-GS on page 10)

21. Page 11, line 40. When the PNNS-GS is being described, is the term “moderation in consumption” referring to “portion sizes”?

The ‘moderation in consumption’ recommendation in the PNNS-GS refers to both portion size and frequency of consumption which combined defined the average daily intake. Depending on the component, the item refers to food groups’ consumption (servings/day) or nutrient intake (g or mg/day). For moderation in consumption, the PNNS-GS refers to nutrient intake: for sugars, it takes into account the percentage of added sugars from sweetened foods in the overall energy intake; for fats, it takes into account the percentage of lipids from added fats in the overall energy intake; for salt it takes into account the total amount of salt in the diet (g/day).

The indicators of the PNNS-GS are reported below. Should the reviewer require it, we are willing to include this table as supplementary material.

Table 1 Components in the modified PNNS-GS

Recommendation Scoring criteria Score

Fruit and vegetables At least 5/day [0-3.5] 0

[3.5-5] 0.5

[5-7.5] 1

≥7.5 2

Bread, cereals, potatoes and legumes At each meal according to appetite [0-1] 0

[1-3] 0.5

[3-6] 1

≥6 0.5

Whole grain food Choose whole grains and whole grain breads more often [0-1/3] 0

[1/3-2/3] 0.5

≥2/3 1

Milk and dairy products 3/day (for ≥55years-old; 3 to 4/day) [0-1] 0

[1-2.5] 0.5

[2.5-3.5] (for ≥55 years old 2.5-4.5) 1

>3.5 (≥55years old >4.5) 0.5

Meat, poultry, seafood and eggs 1 to 2/day 0 0
 [0-1] 0.5
 [1-2] 1
 >2 0.5
 Seafood At least 2/week <2/week 0
 ≥2/week 1
 Added fat Limit consumption Lipids form added fat >16%EI/day 0
 Lipids form added fat ≤16%EI/day 1
 Vegetable added fat Favour fat of vegetable origin No use of vegetable oil or ratio veg oil/total added
 fats ≤0.5 0
 No use of added fats or ratio veg oil/total added fats >0.5 1
 Sweetened foods Limit consumption Added sugar from sweetened foods ≥17.5% EI/day -0.5
 Added sugar from sweetened foods 17.5-12.5% EI/day 0
 Added sugar from sweetened foods <12.5% EI/day 1
 Non-alcoholic beverages Drink water as desired; limit sweetened beverages to ≤1glass/day <1l
 water and >250ml soft drink/day 0
 ≥1l water and >250 ml soft drink/day 0.5
 <1l water and ≤250 ml soft drink/day 0.75
 ≥1l water and ≤250 ml soft drink/day 1
 Alcoholic beverages Women ≤2 glasses of wine/day; Men ≤3 glasses of wine/day Ethanol > 20g/day
 women and >30g men 0
 Ethanol ≤20g/day women and ≤30g men 0.8
 Abstainers and irregular consumers (<1/week) 1
 Salt Limit consumption >12 g/day -0.5
 [10-12] g/day 0
 [8-10] g/day 0.5
 [6-8] g/day 1
 ≤6 g/day 1.5

22. Page 11, line 43: why was adherence to physical activity recommendations not taken into account?

Physical activity in itself was used as a covariable in the main model. Therefore it was not included as a component to the PNNS-GS. Our aim was to use a food-based index of the nutritional quality of the diet. We believe that including both the physical activity component of the score and physical activity would lead to over-adjustment in the model.

23. Page 11, line 53: were energy expenditures based on physical activity data from the IPAQ used? Indeed, energy expenditure included in the PNNS-GS computation (for the penalty computation) took into account IPAQ data.

Results

24. Page 12, line 45: 38,604 participants completed the FOP questionnaire. What is the cooperation rate (e.g., of how many total participants)? Also, it would be valuable to include percentages (e.g., 21,702/38,604).

As suggested, percentages were included in the results section. It now reads:

“Overall, 38,604 subjects completed the questionnaire pertaining to the perception of the various FOP labels formats. Among these, 714 (1.85%) were excluded because they never engaged in grocery shopping. Among the 37,890 remaining subjects, 16,188 (42.72%) were excluded for incomplete data on covariates (the vast majority of which (N=13,066, 80.71% of excluded subjects) for incomplete data on mPNNS-GS computation, which requires the presence of three 24h records, frequency questionnaire on alcohol consumption and frequency of seafood consumption, leading to an overall sample of 21,702 participants for analysis (e.g. 56.22%).”

25. A clear description of the positive and negative dimensions of perception is required. Moreover, how were the dimensions assessed? Is it a composite score?

The mention of 'positive' and 'negative' dimensions was used as descriptive, in order to distinguish between items yielding responses on the preferred vs. less preferred labels. These descriptive dimensions were used to develop the figure, in order to make its interpretation more clear. They therefore do not directly refer to a specific underlying score

26. This is a minor suggestion, but I think it would be best to use a consistent number of decimal places throughout the paper.

As suggested, the number of decimal numbers has been set to 2 throughout the paper for consistency

Discussion

27. Page 15, lines 24 to 27: I am uncertain what this sentence is conveying. "Appealing, through complementary strategies, such as FOP nutrition labels, to those lacking the potential to translate nutritional recommendations or to those to which it would most benefit therefore appears of crucial importance."

To improve clarity, the sentence was modified, as follows:

"Therefore, the fact that the Nutriscore appears to appeal to subjects with low adherence to nutrition recommendations may be a key element to help translating nutritional recommendations into practice, in particular for those with low nutritional knowledge."

28. Page 16, lines 39 to 56: The finding related to colour, and readily interpretative colour-coding, is interesting.

We thank the reviewer for his positive feedback on this finding

29. Based on the weaknesses of this study, including the cross-sectional study design, untested survey measures, and skewed sample, is it possible to conclude that the results of this study suggest the Nutriscore is a legitimate choice for policy-makers in France? I also question whether being more explicit that the authors of this study were involved in developing the Nutriscore system is warranted.

As suggested the conclusion was modified and toned down. Moreover, the mention that the authors developed the Nutriscore is mentioned early in the manuscript

"To conclude, FOP nutrition labels could be useful strategies to tackle social inequalities in nutrition and health, provided that the graphical format that is selected has a wide reach in the population. This is all the more important that subjects who are more concerned about their diet (and more likely to have a healthier diet), are also more likely to use a nutrition label when grocery shopping (20). As such, the Nutriscore, which has a favourable perception among subjects with low adherence to nutritional recommendations, may be a helpful strategy to lead them towards healthier diets."

Tables 2 and 3

Please use consistent headers throughout. For example, does GDA refer to modified reference intakes?

As requested, the headers have been modified and are now consistent across tables. As pointed out by the reviewer, GDA refer to modified reference intakes.

Reference List

1. Haut Conseil de la Santé Publique. Avis relatif à l'information sur la qualité nutritionnelle des produits alimentaires. 2015. Paris, HCSP. 9-11-2015.
2. Lassale C, Galan P, Julia C, Fezeu L, Hercberg S, Kesse-Guyot E. Association between adherence to nutritional guidelines, the metabolic syndrome and adiposity markers in a French adult general population. *Plos One* 2013;8:e76349.
3. Touvier M, Kesse-Guyot E, Mejean C et al. Comparison between an interactive web-based self-administered 24 h dietary record and an interview by a dietitian for large-scale epidemiological studies. *Br J Nutr* 2011;105:1055-64.