Answer to recommender and referees - Stage 2

Removing barriers to plant-based diets: assisting doctors with vegan patients

Romain Espinosa, Thibaut Arpinon, Paco Maginot, Sébastien Demange, and Florimond Peureux

November 9, 2022

In the following pages, the paragraphs are copy-pasted from the reports, our answers are displayed below.

Report #1: Joshua Tasoff

> 2A. Whether the data are able to test the authors' proposed hypotheses (or answer the proposed research question) by passing the approved outcome-neutral criteria, such as absence of floor and ceiling effects or success of positive controls or other quality checks.
> Yes, I believe it does. Furthermore the authors explicitly address floor and ceiling effects in their manuscript.
> 2B. Whether the introduction, rationale and stated hypotheses (where applicable) are the same as the approved Stage 1 submission.
> As far as I can tell the introduction, rationale and stated hypotheses are the same as in Stage 1. Edits are all reasonable and transparent.
> 2C. Whether the authors adhered precisely to the registered study procedures.
> Yes, as far as I can tell this is entirely the case.
> 2D. Where applicable, whether any unregistered exploratory analyses are justified, methodologically sound, and informative.
> The exploratory analysis is descriptive. There are no statistical tests so there is no concern on this point. I find it informative and methodologically sound.

Thank you for the positive comments.

> 2E. Whether the authors' conclusions are justified given the evidence.
> On this point, I have a few minor points. First, I think the authors should cite the nutrition information in the paper when describing the biological tests. The authors say, "Tests T2 to T8 are unnecessary as the patient has no greater deficiency risk than someone who consumes animal-based foods.". I think this needs to be supported.

Thank you for this comment. We have added the following footnote to support the information.

"Alles et al. (2017)[69] report the average daily nutrient intakes of the largest French cohort (Nutrinet) and compare them with the French national nutrition recommendations (ANSES). Without considering supplements, French vegans have on average a daily intake of 760 mg of calcium (recommendation: 750 mg/day), 481 μg of Vitamin B9 per (250μg/day), 10mg of Zinc per day (9.3 mg/day), 18.6 μg of iron per day (6 μg/day), and 248.8 μg of Iodine per day (150μg/day). As far as vitamin D is concerned, the deficiency risks are similar for vegans as for the general population, as sun exposure is the main source, and intakes through nutrition are very limited. However, vegans have a daily intake of 2.7 μg of Vitamin B12 per day, below the recommended intake of 4 μg/day."

The more important point is on this: "In the treatment condition, 21 participants spent less than one and half minutes reading the booklet and were therefore excluded from the following analysis as planned in the pre-registration." But this introduces a selection effect. There is no similar exclusion criterion on the control side. The original exclusion condition in version 1 states, "We will exclude from the data analysis participants from the two conditions who spend less than one and half minutes on the survey." There seems to be an inconsistency here or I'm misunderstanding. The control condition has no booklet and only a survey. The treatment condition has a booklet and a survey. Was the time spent on the booklet and survey recorded together or separately? Does the exclusion criterion apply only to survey time or book+survey time? I think this should be clarified. Ideally, the timing only applies to the survey time and thus the exclusion criteria is applied symmetrically to both conditions.

Thank you for pointing this out. We shouldn't have excluded these observations. We rewrote the entire paper with the new results.. We clearly say that the exclusion concerns the survey itself, not the booklet. No participant must be excluded then. The effect on PMPI becomes "weakly successful or failure". We put our former results (i.e. the analysis with the exclusion rule) in the exploratory analysis.

Report #2: Bence Palfi

Did any prespecified data quality checks, positive controls, or tests of intervention fidelity succeed? - Partly yes. The third pre-registered hypothesis (test of VPI scores) was not tested as the power was not sufficient. - However, the pre-registered outcome neutral tests that meant to check ceiling and floor effects are not reported (only on the pilot data), so this cannot be assessed.

Thank you for this comment. We added the fact that there were observations at the lower and upper bounds of the scales so that we used Tobit estimations.

Are any additional post hoc analyses justified, performed appropriately, and clearly distinguished from the preregistered analyses? Are the conclusions appropriately centered on the outcomes of the preregistered analyses? - The exploratory section introduces post-hoc comparisons, but these are only based on descriptive statistics and not on statistical tests. I agree that these comparisons are interesting, as it may reveal that the information campaign did not improve the understanding of all issues (e.g., Zinc test), but, I think these comparisons need proper post-hoc statistical tests to substantiate any claims made in that section or later. That is if you want to claim that the information campaign improved the understanding of all medical tests but that of the Zinc test, then you need to back this claim up with statistical analyses. - The rest of the exploratory analyses look good

Thank you very much for your feedback. We agree with you that we were actually too fast on this. We rewrote the associated paragraph and now give statistical evidence supporting our claims. It now reads:

> *"Second, the information intervention has however a limited or even null impact on (hypothetical) medical prescriptions. Exploratory analyses suggest that the intervention increases the prescription of T1 and decreases the prescription of several other tests (T2, T3, T5, T6). The intervention seems to have limited or no impact on T7 or on T8. More surprisingly, the intervention seems to increase the prescription rate for the Zinc test (T4). In the control group, 28% of the doctors ask for a Zinc test while 51% of them do so in the treatment group. It suggests that our booklet can be improved regarding the information provided for Zinc. In the experiment, we mentioned that Zinc concentration levels tend to be smaller for vegan patients than for the general population but that there is no significant clinical impact. Doctors might have retained that Zinc levels can be smaller and might have preferred to test for it. When we drop the Zinc variable, the estimated marginal treatment effect on PMPI (scaled between 0 and 1 in both cases) goes from 0.044 to 0.096. In this case, the one-sided p-value decreases to less than 0.001. This suggests that the lack of results is likely to be due to the Zinc effect."*

- I believe that the Discussion section could be expanded, mostly focusing on the results of the confirmatory analyses and their interpretation.

Thank you very much for your comment. We expanded the discussion following your comment and the comments of the other referees.

Are the overall conclusions based on the evidence?
I agree with the conclusions regarding the first hypothesis.
However, regarding the second hypothesis, I have two issues:
- you should bear in mind that you ran a vignette-based study rather than measured actual performance in general practice. Hence, I would refrain from saying that the information campaign improves medical practice, instead, I would say that it has the potential to improve it. This is especially relevant for the abstract, which does not reveal that it is a vignette-based study (perhaps, it would be nice to add this too to the abstract)

Thank you for this comment. We agree that the clarity on this point could be improved. We have modified the abstract accordingly and now mention "*case studies*". We have also modified accordingly in the text when necessary.

- the observed effect size was half of the pre-registered SESOI. I had the impression that this is not well pronounced in the abstract and the discussion section. More attention should be given to this and explained why the observed effect is still interesting (see also my comment later about the lack of information on how the observed effect sizes should be interpreted).

Thank you for this comment. We discuss in the discussion how we can interpret the strength of our effect compared to previous works on the impact of medical information campaigns.

Major comments

Data sharing link (https://github.com/EspinosaRomain/DoctorsVeganDiets) is not accessible to me so I could not verify the data and analyses

We are sorry for this issue, the project was in private access. We have switched it to public access and it should now be available. Please let us know if you still have issues accessing the data and codes.

It was a bit unfortunate that the VPI scores were not analyzed. I think the power analysis of the VPI scores may have used an unfairly large effect size as the smallest effect size of interest. Doubling an observed effect size is likely to produce an overestimation of a realistic effect, so

this analysis was bound not to meet the 80% power criterion. In my opinion, even a 6-percentage point improvement in donations could be worthwhile. I understand that the approved protocol should be followed, but it would be nice to see this analysis in the exploratory section, now that you have collected the data.

Thank you for this comment. You and Alaa Aldoh (referee #3) made the same comment. We were reluctant to expand too much the exploratory analyses but we decided to add the following paragraph:

> *"Fifth, our experiment contained the Veganism Promotion Index (VPI), which aimed at capturing more active behaviors in favor of the promotion of a plant-based diet. We did not analyze the treatment effect on the VPI in the confirmatory analysis because it did not pass our outcome-neutral test (i.e., low statistical power). The lack of statistical power is mostly driven by the very large share of observations at the boundaries of the scale, e.g., doctors who do not give anything to the information campaign (49% in the control group, 51% in the treatment group). The average donation score is the same in both conditions (mean: 38.7% of endowment). Figure 5 shows that the distribution is three-modal with peaks at 0%, 100%, and 50% (in decreasing order of prevalence). While the VPI does not seem to be affected by the treatment, we observe however a strong correlation with the VDI. A Tobit regression indicates that going from 0 to 100% in the VDI score decreases donations by 46.5 percentage points (p<0.001)."*

It would be great if the discussion would connect the findings to the existing literature on information interventions. There are no references in the discussion section now.

Thank you for this suggestion. We have added the following paragraph in the discussion section:

> "*Our intervention seems to have an effect on the average of those observed by other studies on medical information campaigns. Giguère et al. (2020)[66] provide a systematic review of the impact of printed educational materials on medical practices and healthcare outcomes. The authors report a mean Cohen's d of 0.41 across a range of studies (Cohen's d range from 0.04 to 0.79). We find that our booklet outperforms previous studies on the VDI (Cohen's d = 0.71) and performs below average for the PMPI (Cohen's d = 0.22; Cohen's d = 0.35 when Zinc is removed). Due to the heterogeneity of practitioners, practices, and information, it is difficult to compare effect sizes across studies. Nevertheless, our booklet synthesizes information, requires a minimal cost of time and money for doctors, and is tailored to the specific needs of doctors, which seem to be the most important factors to convey information in the medical field.[67,68]*"

I believe that the discussion could also provide an interpretation of the observed effect sizes of the confirmatory analyses. The smallest effect sizes of interest were defined, so it would be great to put the observed effects in context.

Thank you for this suggestion. We have compared our effect sizes in terms of Cohen's D with a systematic review of studies on informational materials in the medical field (see comment above). However, it is difficult to directly compare the effect sizes and the impact of information campaigns across studies as each have specific characteristics and target a particular health outcome.

The analyses presented in the Discussion section should be moved to the exploratory analyses sections, and the results could benefit from more interpretation/link to other findings within this study or by other studies.

Thank you. We moved the analyses of the discussion to the exploratory analyses section as you suggested.

It would be great to expand more on the limitations of the current study (e.g., using vignettes, how results may generalize to French doctors [is the sample truly representative of French doctors?]) and future directions

Thank you for this suggestion. Yes, the two groups of doctors are representative of the French doctor population (ensured by the polling institute). We have added the following paragraph:

> "*The results of the study are positive and show that the booklet has the potential to improve doctors' opinions and, to a smaller extent, medical practices with vegan patients. Our sample is representative so that the results are generalizable to the entire population of French doctors'. Further studies are required to assess the effectiveness of the conveyed information. The study is based on case studies, thus reflecting stated rather than revealed preferences, and may therefore not reflect doctors' actions in real practice. Responding to case studies involves less responsibility for doctors, allowing them to express more sympathetic views and adopt riskier behaviors without consequences. Future studies should focus on the impact of the booklet on real life cases as well as the long-term effect of the information on medical practices.*"

Minor comments:

r 150: "diet" is missing after plant-based

Thank you for noticing, we have added it.

r 177: the control condition is called baseline condition here.

Thank you for noticing, the text now reads "*control vs. treatment*".

There is no reference to Figure 5 in the text.

Thank you for this comment. It is now figure 6. We have added the following at "*Figure 6 reports the VDI recoded answers for the control group.*"

The caption of Figure 5 should clarify that the presented data only includes doctors from the control group.

Thank you for this suggestion. We have added the mention "Control group only (N=200)." in the captions of figure 5, 6 and 7.

Report #3: Bence Palfi

> The study is well executed and the authors have carefully considered my prior feedback. The reporting of the statistical analyses could be clearer: a) report the statistical results fully (not just p-values), and b) report the SD/SEs alongside every mean. I am slightly confused by why the authors conclude that information had a weakly positive effect on PMPI scores when the scores were lower than the SESOI. My understanding is that you would suspend judgment in this case, but it is likely that I've misunderstood.

Thank you for your comment. During the Stage-1 process, we agreed to decide based on the protocol described in Figure 2. If the treatment effect has a non-null effect but this effect is smaller than the SESOI, we committed to say that the treatment is weakly successful. Last, we added in the text the SD and SEs with the average estimates in the confirmatory analysis section. We also added a table in Appendix that comprises all means, SDs, estimates, standard errors and p-values. The table is reported below.

	Control group		**Treated group**		**Marginal effects**		
	Mean	**SD**	**Mean**	**SD**	**Estimate**	**SE**	**p-values**
VDI total	**0.550**	**0.214**	**0.382**	**0.226**	**-0.166**	**0.021**	**<0.001**
VDI 1	0.409	0.328	0.269	0.308	-0.142	0.032	<0.001
VDI 2	0.691	0.273	0.501	0.325	-0.185	0.030	<0.001
VDI 3	0.382	0.326	0.247	0.297	-0.125	0.031	<0.001
VDI 4	0.473	0.330	0.644	0.306	0.174	0.032	<0.001
VDI 5	0.259	0.271	0.464	0.301	0.213	0.029	<0.001
PMPI total	**0.293**	**0.228**	**0.348**	**0.273**	**0.044**	**0.024**	**0.069**
T1	0.060	0.238	0.110	0.315	0.050	0.028	0.073
T2	0.730	0.445	0.645	0.480	-0.085	0.046	0.067
T3	0.800	0.401	0.625	0.485	-0.175	0.044	<0.001
T4	0.280	0.450	0.500	0.501	0.220	0.048	<0.001
T5	0.765	0.425	0.560	0.498	-0.205	0.046	<0.001
T6	0.900	0.301	0.800	0.401	-0.100	0.035	0.005
T7	0.635	0.483	0.590	0.493	-0.045	0.049	0.357
T8	0.605	0.490	0.610	0.489	0.005	0.049	0.919
PMPI no zinc	**0.232**	**0.242**	**0.326**	**0.282**	**0.096**	**0.026**	**<0.001**

Although the study did not have enough power to test effects of the intervention on VPI, I think it is worth including the results of the analysis either in the exploratory analyses section or in the supplementary materials.

Thank you for this comment. You and Alaa Aldoh (referee #3) made the same comment. We were reluctant to expand too much the exploratory analyses but we decided to add the following paragraph:

> *"Fifth, our experiment contained the Veganism Promotion Index (VPI), which aimed at capturing more active behaviors in favor of the promotion of a plant-based diet. We did not analyze the treatment effect on the VPI in the confirmatory analysis because it did not pass our outcome-neutral test (i.e., low statistical power). The lack of statistical power is mostly driven by the very large share of observations at the boundaries of the scale, e.g., doctors who do not give anything to the information campaign (49% in the control group, 51% in the treatment group). The average donation score is the same in both conditions (mean: 38.7% of endowment). Figure 5 shows that the distribution is three-modal with peaks at 0%, 100%, and 50% (in decreasing order of prevalence). While the VPI does not seem to be affected by the treatment, we observe however a strong correlation with the VDI. A Tobit regression indicates that going from 0 to 100% in the VDI score decreases donations by 46.5 percentage points (p<0.001)."*

The report ends somewhat abruptly and I think a conclusion paragraph is needed to tie all the findings together.

Thank you. We substantially expanded the discussion section following your comment and other referees' comments.

Lastly, the analysis code is not available. Otherwise, the results of the study are promising and the authors have done a good job visualizing their data. Well done.

We are sorry for this issue, the project was in private access. We have switched it to public access and it should now be available. Please let us know if you still have issues accessing the data and codes.

Otherwise, the results of the study are promising and the authors have done a good job visualizing their data. Well done.

Thank you for this positive comment!