CRITIQUE OF IIM AHMEDABAD STUDY

by 4 Independent Experts

Comments on the IIM Study : Consumer preference for different nutrition front-of-pack labels in India

Requested by NAPi

By: Dr. K.R.Antony. Public Health Consultant, Kerala. Independent Monitor, National Health Mission, Govt. of India and Former Director, State Health Resource Centre, Chhattisgarh, India.

What is the Objective of the study? Neither in the Executive Summary nor in the detailed Introduction it is clearly stated. We have to figure it out from the final paragraph on the Recommendation.

Why this question is important at the outset, because the methodology should follow meeting those objectives.

If it is "well established that FOPLs have ability to nudge healthy consumption behaviour with regards to packaged foods" as stated in the opening sentence of the Executive Summary then that aspect <u>need not be probed and should be kept out of the methodology.</u>

Is it then narrowing down to "lack of clarity on which kind of FOPL is most comprehensible, acceptable and yet effective" Methodology is to find out which of the five popular FOPLs is "the easiest to understand and influences purchase intentions alike"?

Tests were done on ease of understanding and change in purchase in intentions according to the Executive Summary. Respondents are asked to rate different aspects of FOPLs.

Classification of foods into healthy or unhealthy is a technical or professional step. This should be decided by a Governance body like FSSAI or Food and Nutrition experts or professionals, not even by the Food industry or Manufacturers. So then why do we seek the opinion of the Consumer knowledgeable or illiterate. It is pointless.

What is the role of categorization into three groups treated with "no health prime, healthy prime, non-healthy prime" Is there any added advantage to this extra step? Isn't it confusing the respondent and prevents her or him from expression of the original understanding and opinion? The authors justify its purpose as "to judge the relative effectiveness of the different FOPLs as a signage for healthy and unhealthy foods". The priming of the respondent is an unnecessary step in the methodology. If at all, the label must only carry a healthy prime to take a positive step with regard to decision to purchase. Ideal will be to provide the crystallized information in best acceptable way for consumption and leave it open for consumer to decide. We have done it with Tobacco packs and Alcohol bottles.

The methodology states that profile of the respondent is captured after the respondent has made a choice rating of the FOPL. There is no exclusion and inclusion criteria based on the profile of the respondent, before taking an opinion poll.

65-72% of the respondents are in the habit of reading labels. Automatically 28-35 % of respondents who do not read labels should be excluded from making a relative comparison between labels. This auto exclusion would have given more accurate information. Will you ever ask for opinion on relative merits of three comparable brands of whiskey to a teetotaller?

It is not a wise assumption that all people whom we interview know the scientific basis and concepts of basic Nutrition and they must be on the lookout for a commercial food product that can be beneficial or potentially harmful. Unless they have a responsibility to self-protect, they won't be bothered about warning on packaged foods. There must be some tools to assess this basic understanding level and then decide on inclusion or exclusion of the participant from the study.

The participants of the study must have the capacity to objectively evaluate the FOPL based on the information content, must have the ability to compare, identify least harmful, or identify higher content than recommended and complete the task.

Since the study revolved around only two packaged foods, biscuits and chips and their value share on marketing, purchase and consumption patterns, the preponderance of urban sample over rural across the country and in each state, what I notice, is justified. This is against the common dispersion of sampling units, (population proportionate sampling) in other population-based studies where there is a weightage for rural over urban areas.

Regarding the segment who can and would purchase packaged food, majority are from the affordable groups and only 31.6% of the respondents are from the less than Rs.10,000/- monthly household income. With regard to educational background what is the use of asking preferences on FOPLs to 13.8% of respondents with no schooling at all. In my opinion, they should have been exempted from quizzing.

Why respondents below 18 years are totally excluded? Young adolescents are big consumers of packaged foods like biscuits, chips and bottled soft drinks. There are 17.5% above 18 years who are students among the classification of respondents by Occupational groups. Occupational background hardly makes any difference with regard to purchasing biscuits and chips including 18.4% who are wage labourers. Wage labourers do buy packaged foods with spreading consumeristic culture spreading from urban poor to rural poor. Their decision making is based on the imitation factor or peer pressure, rather than informed choice.

 What is overlooked in the sampling is the fact that the students above 6 years to 18 years who are big consumer segment of packaged unhealthy foods from their pocket money or they pressurise parents to buy. That is significant target population whose decision making with regard to purchase can be modified by scientific information and positive nudging. Importance of this segment is amply highlighted in Comprehensive National Nutrition Survey.

The primes whether healthy or unhealthy worked in reinforcing their opinions and influencing purchase positively or negatively according to the type prime used claims the authors. This only underscores the fact that gullible can be influenced by any maneuverer. Do we need another dimension in this study to re-establish that known fact then and complicate the methodology? The conventional wisdom is to package scientific facts to convince consumer and expect change "decision making" and behaviour positively. Anyway, in the conclusion and recommendation of the study, "Priming" healthy or unhealthy does not feature at all in its influence.

Price, brand, flavour and expiry date are influencing factors in decision making irrespective of Prime factor.

 What do they want on the label? 45 % of control group want health risk related information, 35% wanted Nutrient composition and only 20% concerned about weight gain. Weight gain related, figure conscious, middle class and upper class, urban consumer looks for package information seriously but that is about a fifth of the market share. This is an important finding of the study with Policy implication on the warning labels.

The report does neither attach the survey questionnaire nor display the visual tools used to get the responses.

Overall, this is a very elaborate study in which some avoidable methodological errors have crept in. With exclusion of data responses from certain subsets for reasons explained above, from among the disaggregated data tables, it can make the findings crisper and more realistic. For example, elimination of responses from illiterate, those who never read food labels etc. We cannot do anything about the missing data from young adolescent children from age 10-18 years now and that void will remain.

In conclusion, no firm Policy guideline tips can be derived from the findings of this study.

…………………

Critical appraisal of the FOPL study

PHFI Review by:

Prof. Suparna Ghosh Jerath, Prof. Monika Arora
Dr. Niveditha Devasenapathy (The George Institute for Global Health)
May, 2,2022

Title of the study: Consumer preference for different nutrition front-of-pack labels in India

Summary of the study

a.   **Design:** A large scale individual randomized control trial
b.   **Population:** Community dwellers across age groups (18-60+ years)

c.   **Interventions:** Exposure 5 FOPL with three different variants, another level of stratifications on priming (healthy, unhealthy, no priming)

d.   **Control:** No FOPL, and then 3 levels of stratifications on priming (healthy, unhealthy, no priming)

e.   **Outcome:** Behavioral: a) Purchase intention and b) ease of identification, understanding, reliability, complexity of label, label detecting the presence of unwanted nutrient. The primary outcome variable is purchase intention in the methodology but the results mention the ease of identification, understanding, reliability, and influence as primary outcome variables.

f.   **Timing of outcome measurement:** Right after exposure to FOPL categories

Queries and comments

A. Study protocol

Query

Was this RCT pre-registered with a detailed study protocol and a priori definition of the outcome?

Comment

It will be helpful to know if the trial protocol was registered with sufficient details which would help us to assess any bias due to selective reporting of outcomes.

B. Methodology

B.1 Sampling and recruitment

Queries

The sequence of events

- How was the sample selected and recruitment done?
- How and when was the randomization done? The computer-generated randomization was done on which universe?
- How was the stratification done (prime/age/gender/urban/ rural)? Although the total sample size (20,564) is large enough, is the study powered for so many sub-group analyses?
- Around 62% of interviews were conducted in physical presence, whereas 38% were conducted online- how was randomization followed in an online survey, especially in rural areas?
  Also, how was representativeness ensured as those not using smartphones would have been left out of this study?
- How many contact points were there?

Comments

- The trial protocol should provide detailed information on the sequence of implementation of intervention such as contact, consent process, unit of randomization, implementation of randomization, priming, and exposure to different FOPL. This sequence is not clear in the study methodology.
- Need elaboration on rationale and reference for using prime in the study.
  - **Examples**
    - Sample size has been specified in each prime but there is a need to elaborate upon how these sample sizes were arrived at and how were participants recruited in these primes?
    - After the choices were made, participants were asked to self-report on socio-demographic variables that include gender, age, occupation, etc. Please elaborate, what is the rationale for probing this information after choices have been made? Were any participants excluded after getting this information?
- The study was conducted on 18-60+ years, why have the authors not considered below 18 years to be not included in the study, as this is the age group which is largely impacted. Healthy behaviours and choices get etched at a very young age and the attractiveness of pack, schemes and promotions accompanying products influence the family's decision to buy specific products.

B.2 Ethical considerations

Comment

- The study involves data collection on human subjects but nowhere mentions about ethics
approvals, consent procedures or documentation of the consent. If a waiver of consent was sought, then this should be explicitly mentioned in the report.

B.3 Intervention

Queries

- What were the exact questions asked to the participant?
- What was the language used for asking the questions? How have they defined the 6 items used to assess the effectiveness of types of FOPL in the local language?
- Was the questionnaire used for eliciting information validated?
- What was the language used in the package label?
- How was the intervention standardized? How were the interviewers trained? Since this is a Pan-India study, was the tool translated into different regional languages for better comprehension? Was a standardized protocol used for carrying out the entire data collection process?
- Who administered the intervention? Was (s)he the same person who randomized the participant into a group and then did priming (3 types) and

then showed the label to capture the "decision to purchase" response and then also conducted the interview?

- What was the finding of pre-testing and how were they included in the final questionnaire used for the main study?

Comments

- FOPL labels had one of the six items "Label helps detect the presence of an excess of an unwanted nutrient". This would be influenced by nutrition literacy and perhaps the educational status of the respondents. Was any comparison drawn on baseline knowledge or health awareness score of respondents in prime vs no prime groups? Also, it would be interesting to know how these questions were framed for the diverse participants.
- The questions were pre-tested with 77 participants across states with representation across gender, age group, and education level. What about rural and urban representation at this stage?
- There is a need to address bias when the entire process seems to be "not blinded"
- In a trial, participant characteristics in each arm should be balanced. This validates the robustness of randomization. The report does not present or discuss baseline comparability of the key characteristics. Was there any effort made in this direction?
- Both interviewer and participant could see the label. Since person who enrolls, administers the intervention and measures the outcome is the same person how was performance bias and ascertainment bias addressed in this trial?
- It is a behavioral intervention. Does one-time priming bring about a behavior change? And why was priming given? In an RCT all the previous exposure and knowledge are supposed to balance out across the different arms.

C. Results and outcome

Queries

- Table 1: Different variants for each label type have been specified. Information on these variants needs to be elaborated upon for a better understanding of each treatment arm.
- Table 4: Mean scores and standard deviations for each group are mentioned; however, group differences, 95% CI, and the significance (p) value is not mentioned. Was the mean score for HSR significantly higher than all other types of FOPL?
- Table 5: How were scores converted to ranks? Was this an a priori decision to analyze using ranks?

- Table 6: Is the mean score of the warning label significantly higher than HSR? Was any statistical test applied to prove that ranks of HSR were better than warning labels?

Comments

- We do not see baseline characteristics table of the participants across treatment arms and the control arm (stratification by geography/ age/gender). Were they similar?
- Behaviour to purchase chips or biscuits is clustered in specific geographies. Did the authors adjust for clustering during the analysis?
- There are 15 treatment arms considered at some places, while they are collapsed to 5 treatment arms at others. This was based on the type of FOPLs and the variants were disregarded.
- How were the subpopulations selected to study the relative performance of labels? The findings across gender, age groups, place of living (rural and urban), and occupation are inconclusive. The authors' conclusion of HSR as the choice of FOPL is unconvincing. MTL, HSR and warning labels are also being reported as preferred modes of FOPL in different strata. Stratification should have been done based on age, geography, demography, and education instead of priming.
- Interaction between priming and FOPL needs to be well documented.
- Bias due to selective reporting of outcomes cannot be ruled out.

D. Overall comments

- This study lacks the scientific rigor, ethical requirements, and reporting requirements of an experimental design. The research question and primary outcome variable are not consistent throughout the study document. A time-stamped trial protocol will help the reader to know how much of the conclusion is based on post hoc decisions.
- The methodology is not explicit, unnecessarily complicated and bias due to selective reporting of outcomes in the reported results, and the conclusions cannot be ruled out. A very complicated study design with several levels of stratification with no conclusive findings.
- Very difficult to decipher the findings owing to several subgroups (perhaps not needed)
- A lot of ambiguity in terms of the behavioral intervention given and sequence of events (recruitment, different components of intervention given, and data collection)
- Geographic and demographic representation is well calculated. However, how this strategy was implemented is not well explained. What was the universe?
- There is no clarity on the statistical analysis undertaken at different stages of the study.

- The study appears to be perception-based marketing research and perhaps not an epidemiology-based RCT.
- The study needs to be re-done with all the critical considerations of RCT and the research questions need to be prioritized based on the questions asked from the policy perspective. In fact, this study can be a basis for a large well designed representative randomized controlled trial taking in all the lessons learned from the implementation of the study, intervention, and measurement of outcomes.
- In its current state, the study findings do not meaningfully inform policy on "consumer preference on front-of-pack nutrition labels" in the context of their health and wellbeing.

end of the document

…………………..

Comments on the IIM Study : Consumer preference for different nutrition front-of-pack labels in India

Requested by NAPi

May, 2nd 2022

By: Prof. Dr. Piyush Gupta, MD, FAMS *Professor and Head, Department of Pediatrics, University, College of Medical Sciences, Delhi; Immediate Past President, Indian Academy of Pediatrics*

At the behest of Food Safety and Standards Authority of India (FSSAI), Indian Institute of Management (IIM), Ahmedabad conducted a randomized controlled trial on more than 20,000 consumers to determine their preferences for different FoPLs, in order to frame policy for the Indian Market. The authors recommended that the Health Star Rating to be considered as a preferred choice for policy.

Here are my comments:

1. **Methodology**: For the results and conclusions to be valid, it is important to have a robust unquestionable methodology. More so when this has to be translated into policy and action at the National level by the regulatory body (read FSSAI). Let us therefore examine the applicability of this study for 1.3 billion consumers of this country. The IIM trial enrolled a nationally representative sample of 20,564 face to face respondents who were randomized to one of the six groups: no FoPL, HSR, Nutri-score, Warning labels, traffic lights, and monochrome GDA. The participants in the five intervention groups were further subcategorized into three groups: no health prime group, a healthy prime group, and an unhealthy prime group. Thus a

total of 15 treatment groups were created. Three variants of the label were used but there is no information why. Other than this, the control group also had three subgroups. Interventions were conducted in an heterogeneous manner (62% physically and the rest 7811 over video calls). It is also not clear how blinding was done.

2. **The missing piece**: The IIM study was limited to identify purchase intention for packaged biscuits and chips only. A very important group that was left out was sugar -sweetened beverages (SSBs).

3. **Objective vs subjective assessment:** The decisions were not based on objective measures that would really gauge the ability of the consumer to understand the nutrition information. Objectivity is an extremely important criteria while assessing FoPL. Subjective assessment has multiple confounders and thus is unsuitable for a policy decision with implications for the entire country. Similar studies from Chile, Uruguay, and Mexico are also based on objectivity.

4. **Contd.** As also stated above, the results appear to be based on self-reported measures of understanding, which are by and large, subjective. The questions fail to assess whether the participants are actually able to understand the FoPL. Consumers' decision needs to be based on their understanding of the nutrition content/information and not only on how attractive is a particular label. This is the underlying science behind making this policy.

5. **What was the exact stimulus**?: The IIM report also falters on several other aspects. The report also did not include the exact stimuli (treatment) the consumers were exposed to. Exact questions asked to the participants are not elaborated. Justification is not provided for using the healthy/unhealthy primes. The results of FoPL in the no prime condition were all that were needed, to answer the research question.

6. **The other study**: A study led by the scientists of International Institute for Population Sciences IIPS (2022) among 2,689 adults across 6 states (Assam, Delhi, Gujarat, Odisha, Karnataka, UP) has documented that all 5 FoPL were effective in influencing product perceptions and label reactions. And of these, warning labels showed the largest effect. HSR doesn't let the consumer realize the associated health risk with a particular food/beverage.

7. **Authors Selection:** According to the report, both HSR and warning labels appeared to be the most easily identified, most easy to comprehend, considered most reliable, and able to influence the consumer as compared to the rest of the labels. The report indicated that the Multiple Traffic Light (MTL) was the most preferred for health information and the presence of unwanted nutrients.

8. **Contd.** But the study concluded that HSR appears most acceptable. However, their final word on the best FoPL is to be taken with a pinch of salt as the evidence and justification given in this Report needs much to be desired.

9. If our main concern is to alert the consumer to unhealthy nutrients in terms of quantity or quality, then warning labels consistently score better and show a maximum effect size. Studies are available to show that warning label has the largest effect in the ability of consumer to correctly identify that a particular food has excess amounts of nutrients of concerns. In fact, the IIM report itself acknowledges that scientific evidence exists to prove that warning labels deter people from buying unhealthy food for their child, yet the HSR system has been recommended.

10. **CONCLUSION**: It is surprising why Warning label is left out in favour of HSR. Based on the above comments, the report does not sound valid for use in making a public health policy. Let me also provide my opinion below why India should go for warning labels on unhealthy, ultra-processed food products.

11. **Opinion:** Visually, the presence of one or two stars runs the danger of overcoming the absence of four or there stars. Add to it the fact that we've been taught to use 'stars' in an overt and explicit positive connotation, and it effectively renders the health-star rating system redundant as to alert someone to reject for consumption. The health-star rating is also open to industry-abuse. A health-star rating, in essence, could be interpreted for just about anything. It is well known that companies can benefit from misrepresenting their products at the cost of the consumer's health, many will choose to do so. The inefficacy of the Health-Star Rating System is evidenced also by the willingness of the food industry to adopt it, which views it as the a system that least affects its bottom-line. The whole point of advocating for a Nutritional Warning System as an FoPL is to allow the consumer to make an informed decision in a matter of seconds. FoPL only does that, whichever is the product if it crossed the limits. The health-star rating takes out all pertinent information, leaving the decision-making to emotion and guesswork. The Nutrient Warning Label is a direct, informative alternative to it. It gives the consumer the exact information they'd need without looking at the detailed list of nutrients on the back of the pack. A high amount of sugar, salt or fat can be immediately conveyed without any ambiguity. This is also the reason that the food industry, plagued by inertia, is vehemently opposed to it – it would mean shrugging off the complacency and actually reworking their products to be healthier and less harmful, in order to avoid the label. The Nutrient Warning Label thus has the potential to affect not only the consumer's psyche, but industry practice as well.

………………

Review of the study: Consumer Preferences for Different Nutrition Front-of-Pack Labels in India conducted by IIM-A, Dexter. Feb 15, 2022

Requested by : Nutrition Advocacy in Public Interrst(NAPi)

By Prof. Dr. Abhaya Indrayan,

MSc, MS, PhD (Ohio State), FSMS, FAMS, FRSS, FASc. Former Professor of Biostatistics, University College of Medical Sciences Delhi.

The report describes the methodology and the results of their nationwide survey of 20,564 persons of age 18 years and above with nearly equal representation of males and females. The objective was to assess the preference for front-of-pack labels on food products (the survey was limited to chips and biscuits). These subjects were randomly allocated to 45 'treatment' groups (15 label types x 3 Primes) @ nearly 400 per group. In addition, 800 per Prime were in the Control group with no label. One of the Primes was 'None' and the other two were 'Healthy' and 'Unhealthy'. There are the prompts provided to the respondents regarding the food products. The survey was limited to 5 label types (Multiple Traffic Lights (MTL), Monochrome GDA, Nutri-Score, Warning Labels, and Health Star Rating (HSR). Three variants of each type of labels were also considered but the findings are stated after collapsing them. Thus, the need to have 3 variants of each type of label is not clear.(p 14)

Rationale of determining 400 as adequate sample size for each 'treatment' is not mentioned. These subjects have been divided into States, rural-urban, male-female, age-groups, and occupations – thus the spread of the representation of various categories looks very this. The results are rightly presented after collapsing these categories except at some places. However, the details of the method of random allocation are not provided – thus the findings cannot be taken on their face value. In addition, the sampling frame for the selection of the subjects is not clear, although the allocation of the sample to rural and urban segments is assiduously explained. Nearly one-sixth of the sample was discarded due to incompleteness or for other reasons. (p-17) This is substantial and may have introduced bias. This possibility has not been considered. About 62% interviews were physical and the remaining on video calls – both were rightly pooled because the buying intensions of these two groups were not found statistically significant despite not so small sample.

The rating for preference of the labels and purchase intension were elicited on a 7-point Likert scale. The report concludes that HSR was the most acceptable label, closely followed by Warning Label. However, the report says that MTL was the most preferred for indicating health information and the presence of unwanted nutrients, as well as for indicating purchase intension. Thus, the results were not as unequivocal as the report seems to convey in their conclusion.

From the consumer viewpoint, nutrient information is indeed useful, but perhaps equally useful is the information on the ill-effects of the constituents of the food products. The survey seems to be missing this crucial aspect. Perhaps Warning Labels serve this purpose well and they have been found to get nearly the same rating on average as the HSR. Warning Labels have received a significantly higher rating by 'No Prime' respondents for both chips and biscuits than the 'Healthy Prime' respondents (Table 3). Strongest effect of priming was observed with Warning Labels in the 'Healthy Prime' and 'Unhealthy Prime' groups (p.21). They were also marginally ahead in term of reducing purchase intention in both 'Healthy Prime' and 'Unhealthy Prime' groups (p.22). Warning labels were also observed to have more

extensive support across occupations (p.26). These findings have not received the prominence in the report they deserve.

Statistically, as already mentioned, the adequacy of the sample size is not explained, the sampling frame from which the sample was selected is not specified, and the method of selection is not fully explained. Possible effect of huge nonresponse on the findings is not discussed.

There are other minor statistical problems. The degrees of freedom (dfs) shown for the *t*-tests differ from each other and what they apparently ought to be, and no explanation for this discrepancy is provided (p.17). The *t*-values and *P*-values are shown at some places and not at other places though significance is concluded (e.g., top of p.25). The symbol *d* has been used without specifying that it (probably?) is Cohen's *d*. A complete table with means and *P*-values for the findings mentioned on page 25 would have enhanced the credibility gof the findings. The findings stated as bullet points on page 26 may not be statistically significant (no *P*-values given) because of small *n* in various subgroups. In the absence of a table and the data, the findings on page 27 (particularly the top paragraph) look subjective and unsubstantiated.

Considering all the above-mentioned points, the results and conclusion of the report are suspect. Warning Labels seem to deserve a better consideration.

Conflicts of Interest : None

ends.