

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)	Ultra-processed foods and added sugars in the US diet: evidence from a nationally representative cross-sectional study
AUTHORS	Martinez Steele, Euridice; Baraldi, Larissa; Louzada, Maria Laura; Moubarac, Jean-Claude; Mozaffarian, Dariush; Monteiro, Carlos

VERSION 1 - REVIEW

REVIEWER	Robert H. Lustig, MD, MSL UCSF, Division of Pediatric Endocrinology, San Francisco, CA, USA
REVIEW RETURNED	23-Sep-2015

GENERAL COMMENTS	This is a well documented and designed study which addresses a major question about the contribution of ultra-processed foods to the intake of added sugars in the USA. The results of this study are generalizable as they used the nationally representative NHANES dataset. One consistent issue with studies of this nature is that there is major potential for recall bias since the data was collected by dietary recall. It is also very well known in the literature that people with obesity tend to underestimate their dietary intake. Thus, this might lead to underestimation of the authors' results. The authors also classified the food consumed in the lower processed food group if there was a question about the ingredients, in an effort to minimize the bias and overestimation (as mentioned in the supplemental material). There are several points in the manuscript that requires clarification for the reader. 1- Page 8, line 49-51 "The first dietary recall interview was collected in-person in the mobile examination center (MEC) while the second was collected by telephone 3 to 10 days later." --- The second time around participants might recall better since they were already exposed to the applied questionnaire earlier. Or even perhaps not report true consumption for the same exact reason. Did the authors validate each assessment for validity? 2- Page 8, line 54-56 " Among the 13,272 people screened in NHANES 2009-2010, 10,537 (79.4%) participated in the household interview and 10253 (77.3%) also participated in the MEC health examination(21). --- Was the questionnaire also applied during the household interview? or only in the MEC, and then followed up by phone for a second recall? 3- Page 9, line 58-60 " We evaluated 9,317 survey participants aged 1 year and above who had one day 24-hour dietary recall data and had not been breast-fed on either of the two days." --- Up until what
-------------------------	---

	age the caregiver was the source of dietary recall? Was there a single caregiver for the children to provide dietary recall information? What about babysitters or daycare programs? When the investigators followed up with the phone call for the second recall, did they get in touch with the same caregiver? 4- page 9, line 60 " These individuals had similar socio-demographic characteristics to the full sample of 10,109 participants interviewed." What socio-demographic characteristics: Age? Gender? Income? Ethnicity? Education? What was different? 5- Page 11, line 130: Did you consider level of education as a covariate? 6- Page 13, line 153: " The most common ultra-processed foods in term of energy contribution were breads; soft drinks, fruit drinks and milk drinks.... " --- By "milk drinks" do you mean drinks that are prepared with milk? Milkshakes? Smoothies? Coffee? 7-Page 19, line 204-207 " The strength of the association remained fairly the same after adjusting for the proportion of added sugars in non-ultra-processed energy intake (coefficient for linear term=0.19, 95% CI: 0.17 to 0.23) and for age, sex, race/ethnicity, and family income (coefficient for linear term=0.22, 95% CI: 0.18 to 0.26." --- Did you adjust for age, sex, race/ethnicity and family income in addition to proportion of added sugars? 8- What was different in terms of the characteristics of the subjects in the first quintile versus the rest of the quintiles? What socio-demographic characteristics changed as consumption increases through quintiles? This might yield very telling information.
--	--

VERSION 1 – AUTHOR RESPONSE

1- Page 8, line 49-51 "The first dietary recall interview was collected in-person in the mobile examination center (MEC) while the second was collected by telephone 3 to 10 days later." --- The second time around participants might recall better since they were already exposed to the applied questionnaire earlier. Or even perhaps not report true consumption for the same exact reason. Did the authors validate each assessment for validity?

Analyses using exclusively the first dietary recall interview yielded very similar results to those obtained using two dietary recalls. Nonetheless, in order to avoid the above-mentioned biases and following the second reviewer's advice, we have opted to present results using exclusively the first dietary recall interview.

2- Page 8, line 54-56 " Among the 13,272 people screened in NHANES 2009-2010, 10,537 (79.4%) participated in the household interview and 10253 (77.3%) also participated in the MEC health examination(21). --- Was the questionnaire also applied during the household interview? or only in the MEC, and then followed up by phone for a second recall?

This has been clarified in the manuscript as follows:

The survey included an interview conducted in the home and a subsequent health examination performed at a mobile examination center (MEC). All NHANES examinees were eligible for two 24-hour dietary recall interviews. The first dietary recall interview was collected in-person in the MEC(15) while the second was collected by telephone 3 to 10 days later but never on the same day of the week as the MEC interview(16).

3- Page 9, line 58-60 "We evaluated 9,317 survey participants aged 1 year and above who had one day 24-hour dietary recall data and had not been breast-fed on either of the two days." --- Up until what age the caregiver was the source of dietary recall? Was there a single caregiver for the children to provide dietary recall information? What about babysitters or daycare programs? When the investigators followed up with the phone call for the second recall, did they get in touch with the same caregiver?

This has been clarified in the manuscript as follows:

Dietary interviews were conducted by trained interviewers using the validated (17-19) US Department of Agriculture Automated Multiple-Pass Method (AMPM)(20). For children under 9 years of age, the interview was conducted with a proxy; for children between 6 and 8 years of age, in the presence of the child. Children 9 to 11 years old provided their own data assisted by an adult household member (assistant). The preferred proxy/assistant was the most knowledgeable person about the child's consumption the day before the interview. If the child had more than one caregiver, several individuals could contribute to the Intake data(15; 16).

4- page 9, line 60 " These individuals had similar socio-demographic characteristics to the full sample of 10,109 participants interviewed." What socio-demographic characteristics: Age? Gender? Income? Ethnicity? Education? What was different?

This has now been clarified in the manuscript as follows:

We evaluated 9,317 survey participants aged 1 year and above who had one day 24-hour dietary recall data and had not been breast-fed on either of the two days. These individuals had similar socio-demographic characteristics (gender, age, race/ ethnicity, family income and educational attainment) to the full sample of 10,109 participants interviewed.

5- Page 11, line 130: Did you consider level of education as a covariate?

We have now included educational attainment as a covariate:

All regression models were adjusted for age (1-5 years, 6-11 years, 12-19 years, 20–39 years, 40–59 years, 60 + years), sex, race/ethnicity (Mexican-American, Other Hispanic, Non-Hispanic White, Non-Hispanic Black, Other Race including Multi-Racial), ratio of family income to poverty (categorized based on Supplemental Nutrition Assistance Program (SNAP) eligibility as 0.00–1.30, >1.30–3.50, and >3.50 and above)(14) and educational attainment of respondents for participants aged 20 + years and of household reference person otherwise (<12, 12 years and >12 years).

6- Page 13, line 153: " The most common ultra-processed foods in term of energy contribution were breads; soft drinks, fruit drinks and milk drinks.... " --- By "milk drinks" do you mean drinks that are prepared with milk? Milkshakes? Smoothies? Coffee?

"Milk drinks" refer to drinks manufactured by the food industry prepared with milk. Some examples are provided under footnote 9 in Table 1 ("Including flavored yogurt sweetened with sugar or with low-calorie sweetener, milk-shake, soymilk"). In order to make this clearer we have renamed "Milk drinks" as "Milk-based drinks".

7- Page 19, line 204-207 " The strength of the association remained fairly the same after adjusting for the proportion of added sugars in non-ultra-processed energy

intake (coefficient for linear term=0.19, 95% CI: 0.17 to 0.23) and for age, sex, race/ethnicity, and family income (coefficient for linear term=0.22, 95% CI: 0.18 to 0.26." --- Did you adjust for age, sex, race/ethnicity and family income in addition to proportion of added sugars?

The regression coefficient after adjusting for all 5 variables remained fairly the same. We have now included in the manuscript the coefficient after adjusting for age, sex, race/ethnicity, family income, education and proportion of added sugars in non-ultra-processed energy intake.

The manuscript was been reviewed as follows:

The strength of the association remained fairly the same after adjusting for age, sex, race/ethnicity,

family income, educational attainment and proportion of added sugars in non-ultra-processed energy intake (coefficient for linear term=0.19, 95% CI: 0.17 to 0.22).

8- What was different in terms of the characteristics of the subjects in the first quintile versus the rest of the quintiles? What socio-demographic characteristics changed as consumption increases through quintiles? This might yield very telling information.
We are considering doing this in another article.

VERSION 2 – REVIEW

REVIEWER	Robert Lustig, MD, MSL Professor of Clinical Pediatric Endocrinology, University of California San Francisco
REVIEW RETURNED	08-Nov-2015
GENERAL COMMENTS	Manuscript improved significantly. The comments of the second reviewer and myself were taken into consideration by authors appropriately.