Reviewer's report

Title: Where an individual shops for food and the association with dietary intake, regardless of daily travel patterns among Fayette County adults, Kentucky 2011

Version: 1 Date: 24 October 2012

Reviewer: Angela Liese

Reviewer's report:

Based on data from 121 residents in Fayette County in Kentucky, the authors examined the associations between individual dietary intake and food environment measures (food venue choice and availability of healthy food within the food venue), and the association between food activity space and venue choice and food availability. The attempt at characterizing real shopping behaviors and food availability in stores is commendable and will be a contribution to the literature. My comments and concerns about this study are listed below.

Major Compulsory Revisions:

1. The authors list three hypotheses in the introduction. While the first two are understandable and clear, the third hypothesis sounds circular: “The third hypothesis was that individuals with a healthier food activity space, that is, residents who operate within a space with more healthy food venues relative to unhealthy food venues, would report shopping in healthier food venues, compared to residents who operate within space with fewer healthy food venues.” To my understanding, the same variables inform both the definition of a “healthier activity space” (having more healthy food outlets) and “shopping in healthier food venues” (shopping at healthy food outlets). Please clarify your position or edit accordingly.

2. No background information was found for food activity space in the introduction section. I had also hoped to see some introduction to the theme of shopping behavior.

3. The authors used GPS to measure the “food activity space” in this study. However, how are they determining if this space includes any food shopping? To my understanding, it is the space for “working” for employed individuals or “traveling” for others. Using this method to define food activity space may cause problems if shopping did not occur in the 3 day interval.

4. The classification of food outlets to define RFEI is not clear to me. In CDC’s Children’s Food Environment State Indicator Report (http://www.cdc.gov/obesity/downloads/childrensfoodenvironment.pdf), specialty stores and farmer’s markets are not mentioned in the healthy store category. Also, why are supercenters categorized along with other “unhealthy” outlets?
Lastly, given the ratio nature of the RFEI (and assuming their formula equates to what is described on line 114-115), how did the authors deal with neighborhoods in which no healthy food outlets existed (i.e. numerator equal to 0) or in which no unhealthy food outlets existed (i.e. denominator equal to 0).

5. Line 130-122, I think the assertion that “In the ratio measure used here, higher scores indicated a food activity space with more venues selling more low-calorie healthy food items relative to stores selling energy dense foods.” is not tenable as written, because it is well know that supermarkets carry a large amount of unhealthy products, and in terms of volume, that amount will exceed the amount available in a convenience store. It would be good if the authors would review and revise their language on this matter to present a more nuanced understanding of the issues.

6. The response rate is very low (153/1400=11%) and the demographic characteristics of the sample suggest that this is a middle-to-old sample with high SES. There are about 76% Whites in this county based on Census data, however, in the sample, there are 94% Whites. This issue needs to be addressed as a limitation, as the results can not be generalized to the residents in the county even to those with high SES.

7. In Table 1, are the specialty markets and farmer’s markets combined as one category for type of store for primary shopping? If not, does this mean that not a single participant reported using a farmer’s market as the primary shopping choice? From the percentage, we see that only 10% reported supercenters and 10% reported specialty markets. In the logistic regression in Table 2, when comparing to supermarkets, is the sample size for those reporting specialty and farmer’s market too small? The sparse data may be the reason that large confidence intervals for fruit & vegetable intake for specialty and farmer’s market.

8. With a sample size of 121, a test of interaction is an overly ambitious and statistically risky undertaking and I think the authors would be wise to remove any reference to this effort.

9. The analyses underlying data presented in Table 3 can not be easily connected with the results section pertaining to the same table. I am not sure if it is a matter of inconsistent labeling of if there are actually other analyses being described. In part, the quite confusing layout of Table 3 is making the interpretation very difficult. I would also suggest separating this table into two.

10. At a more conceptual level, the authors need to explain their rationale for presenting their main findings in Table 2 (focusing on primary food shopping venue type) adjusting from the start for both the active living space and the RFEI attribute of the neighborhood. My suggestion would be to start the modeling process without either of the other two variables, because they would be hypothesized to be influencing the shopping behavior and may, or may not, have direct effects on fruit and vegetable intake.
Minor Essential Revisions:

1. In the abstract, the authors mentioned the GPS was worn for a week but in the methods section it was said to be worn for three days. Please clarify and also discuss the implications on the assessment of food shopping behaviors and space. The dependent variables were coded into dichotomous variables for some analysis in the study. How were the cutoffs chosen? For example, for F&V intake 2 times/day was used, for milk 1 time/day etc.. Have the investigators explored the effect of choosing different cutpoint on the results of the analyses?

2. In the logistic models in Table 3, the food venue availability scores were coded as dichotomous variables. Please give some rationale and potential references for the cutoffs, such as 56 for total score and 45 for availability score?

3. The SDs are very small for several variables included in Table 1, e.g. age, years in Lexington and neighborhood. Please verify that these are in fact correct, they seem somewhat implausible, especially for age.

4. What do the p-values mean in Table 1?

5. The initial samples (N=1400) were from a survey on cancer control. Please provide a little more detail on this sample’s characteristics. Are the participants healthy or patients with cancer?

6. In footnotes in Table 2, the abbreviation of CI was provided, but I can’t find a “CI” in the table labeling.

7. Table 2 heading and labeling is rather incomplete. Please spell out the outcomes (e.g. Fruit & Vegetable intake of > X servings/day), improve the titles of the exposure variables (e.g. Primary food shopping venues …, …, …). Furthermore, the convention is the indicate the reference group with 1.0

Discretionary Revisions:

1. The authors mentioned that they asked shopping behavior questions during the interviews. What shopping behavior questions do they include? In the method section, only “food shopping stores” was asked during the interview. Including more shopping behavior questions may enhance the analysis between food environment and dietary intake in this study.

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

'I declare that I have no competing interests'