Reviewer's report

Title: Irregular breakfast eating is negatively associated with adolescents health status and health promoting behavior - a descriptive study in Taiwan

Version: 1 Date: 7 June 2006

Reviewer: Douglas Thompson

Reviewer's report:

General

This article had several noteworthy strengths, especially: 1) the sample was large (n=1617), providing excellent statistical power to detect associations; 2) information about a Taiwanese sample is valuable, given that most studies of this topic have been done in other countries; 3) the authors made laudable efforts to validate their measure of breakfast consumption, including cognitive testing (i.e., discussing the test items with respondents); and 4) associations between breakfast consumption and a wide variety of outcome measures were examined. Because of these strengths, I feel that this article is worthy of publication, provided that the authors are able to address the problems discussed below. As is always the case in reviews, I will focus on problems with the article, but that should not overshadow the fact that this is a valuable article with the potential to make a significant contribution to the literature. None of the issues pointed out here are fatal flaws; all can be addressed without excessive difficulty.

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

Primary issues:

1. The sample is described in a way that is confusing and potentially a bit misleading. Both the Abstract and Discussion state that a “random” sample was used, but the description in Methods indicates that the sample was not random. A random sample (or specifically, a simple random sample) implies that each element in the population had an equal probability of selection. This was not the case in the present study - it appears that students in 10 schools had a high probability of selection, while students in all other Taiwanese schools had zero probability of selection. I don’t want to get into the technical details too much here; the general point is that the sample needs to be clearly and consistently described throughout the article. The sample is in fact a cluster sample (as the authors indicate at one point in Methods). A clear description should involve the following: 1) state how the schools were selected; 2) state how students were selected within schools; 3) state that this is a multistage or cluster sample; and 4) do not describe it as a “random” sample, which it is not. If the authors wish, they could refer to a survey sampling text such as L Kish (1965), Survey Sampling, published by Wiley.

2. Related to point 1, because this is not a simple random sample, it should not be analyzed as such. The participants should not be treated as independent observations; students within a given school are likely to be more similar to each other than are students in different schools. The statistical methodology needs to take account of this non-independence within schools. There are various methods for doing this. Perhaps the most common method for handling students-within-schools designs is hierarchical linear modeling (also known as mixed or random effects modeling); a basic text is Bryk & Raudenbush (1992). Hierarchical Linear Models, published by Sage (I think a second edition is also now available). Another common approach is to adjust standard errors for design effects (basically, variance inflation or deflation), as is done in software packages such as SUDAAN. The software that the authors used for this analysis, SPSS, certainly includes procedures for mixed or hierarchical linear modeling, and perhaps design effects approaches as well (sometimes called “survey” approaches). Regardless of the specific approach that the authors use to deal with this issue, the general point is that the statistical analysis needs to take into account the fact that students within schools cannot be treated as independent observations.

3. The authors appear to have converted all measures (including breakfast consumption as well as all of the outcome measures) to binary variables prior to analysis. It was unclear why this was done, given that most of the variables were collected on continuous or ordinal scales. When variables are dichotomized prior to analysis, information and statistical power are unnecessarily lost. For a basic reference on this issue, the
authors might refer to F Harrell (2001), Regression Modeling Strategies, published by Springer Verlag. The only variable for which I can see a good rationale for converting to binary is overweight, because BMI has often been categorized in this fashion in past articles, thus such a categorization could facilitate comparison with past literature. To address this issue, the authors should either provide a convincing rationale for dichotomizing the variables prior to analysis, or they should analyze the variables on the scales on which they were originally collected (i.e., continuous or ordinal).

4. The statistical methodology needs to be more clearly described in the Data Analysis section. Specifically, it would be helpful if the authors would provide the following information: 1) In Data Analysis, the authors say that they used chi-square tests, yet the table footnotes indicate that Fishers exact test was used; this is confusing. Fishers exact test and chi-square tests may be used to test the same hypotheses, but they are not the same test and they can yield different results. The authors should clearly and consistently describe the tests that they used. 2) In Data Analysis, the authors state that they used multivariate analyses but it is not clear what this means (I guess it probably refers to a regression technique such as logistic regression, or perhaps Mantel Haenzel chi square tests); the authors need to clearly state what techniques were used. Multivariate analysis is not sufficient. 3) In some analyses, the authors adjusted for potential confounds, such as gender, age and household composition. Such statistical adjustment is an excellent thing to do and I applaud the authors for doing it. However, the Data Analysis section should clearly state which variables were adjusted for in which analyses, and why. Otherwise, in Results, it can be unclear whether the results reported are adjusted for these potentially confounding factors, or not adjusted. Further confusion is created because some of the analyses were adjusted and others were unadjusted. One way to reduce this confusion would be to do all statistical tests with statistical adjustment for potential confounds. Unadjusted frequencies can be provided in Tables (as the authors have done) and the unadjusted results do not need to be discussed further in Results, possibly with a few exceptions (for example, I think the authors are correct to report, descriptively, the percent who did not regularly eat breakfast, unadjusted for other variables).

5. Generally I think the article is well-written the length is about right and the organization is clear and logical. However, throughout the article, there are grammatical errors which are distracting and make the article more difficult to follow. I am not going to list all of these; to give a couple of examples, in the second paragraph under Background, poor school performance should be poorer school performance and qualities of life should be equality of life. The article would benefit from a thorough review for English grammar mistakes.

6. The following comments are applicable to all tables. 1) Either in the table titles or somewhere else, the contents of the table cells need to be better described. For example, it is not immediately apparent that column percentages are provided in parentheses; readers should be explicitly told this, they should not need to figure it out. 2) I think row percentages would be more informative than column percentages. To illustrate, in Table 1, for Grade, Middle school would show 73.7% (764/1036) are RBE and 26.3% (272/1036) are IRBE, while for High school 80.9% are RBE and 19.1% are IRBE. 3) Check the column labels in the tables, it looks like RBE is much more common than IRBE, but the text reports that the majority of subjects were IRBE. 4) As alluded to above, it is odd to report a chi-square result in the table, then give a p-value for a different test (Fishers exact test). 5) In a table footnote, give a specific example of how an odds ratio is interpreted. One can figure it from the numbers in the tables, but readers shouldn™t be asked to do so much work.

---------------------------------------------------------------

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

1. For each subscale of the Adolescent Health Promotion scale, it would be valuable to report some descriptive statistics for the present sample, for example mean, median and standard deviation. This could be done in the text or in a table. This would facilitate comparisons with future research using other samples.

2. In Background, the authors state Our ancient forbears acknowledged this lifestyle thousands of years ago (presumably referring to breakfast consumption). If the authors are going to make this statement, they need to be more specific. Although I™m not an expert in the history of breakfast consumption, I™m certainly not convinced that people in the ancient past ate breakfast more often than people do today. If some ancient text does mention the importance of eating breakfast, that would be extremely interesting and the authors should mention it. Otherwise, the authors probably should drop the reference to our ancient forbears.
3. In the first paragraph of Results, there is some redundancy with information presented in the tables. The length of this paragraph could be reduced considerably— it is not necessary to repeat so much of the information that is provided in the tables.

4. In the first paragraph of Discussion, the findings on body weight are not mentioned as a main finding. In fact, I believe that the association between breakfast consumption and absence of overweight should be considered a main finding, because this replicates a result found in several past studies and contributes to a growing set of findings on associations between breakfast and weight control.

---

Discretionary Revisions (which the author can choose to ignore)

**What next?:** Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

**Level of interest:** An article of importance in its field

**Quality of written English:** Needs some language corrections before being published

**Statistical review:** No

**Declaration of competing interests:**

Within the past 5 years, I have received grant funding from General Mills Corp., which sells breakfast foods and thus could receive beneficial publicity from favorable reports about breakfast. However, I do not believe that anything in my review was influenced by my past relationship with this funding source.