

Reviewer Comments:

Reviewer #1: This is a very nice article with important information.

Several inconsistencies or confusing areas should be corrected prior to possible publication. Page 4 line 20 says 2003 BFRSS data files were downloaded. This is misleading. Wasn't 2003 the year of the data used?

Inclusion of specific hypothesis would help clarify the results. The author says primary hypothesis then page 6 says question, then page 7 says second aim.

In the tables indicating which states had significantly different rates of screening between DM and non-DM would strengthen the results. It would also be nice to have overall rates in a category. Page 6 line 10 needs further explanation or perhaps a figure representation of the variables with in the PRECEDE model would help.

Page 7 talks about if the hypothesis 1 was true but this is not discussed previously. If the point of the second aim is to look at factors predictive of screening in woman with DM why was hypothesis 1 important at all?

The discussion and conclusions are very nice and informative. Certainly this is valuable work that brings to light the differences in screening for more than just people with diabetes. Screening needs for un- or under insured women, and women with less education or low income are imperative.

Reviewer #2: This is an interesting article that contributes to the field of health promotion and disease prevention. The inclusion of the PRECEDE model as predictive factors provides theoretical grounding that helps readers understand the significance of the findings.

My comments are largely for clarity as the manuscript is well-written and understandable overall. Data are plural, not singular.

Please rename the Data Analysis to Methods and include a description of your statistical methods used. Presently, that is scattered throughout the paper.

Please expound on the explanation of the contingency table as you have presented it on page 5.

Was chi square used as the statistical test of significance, because odds ratio are reported in the results section and in the tables.

It states that only Missouri and Hawaii do not show a significant association between pap screening and diabetes. However in Table 2, only 1 point separates the women in South Dakota who have diabetes and do not have diabetes and received pap screening. Did the 1 point separation make a significant difference?

The Discussion section includes information about screening rates, but it is not clear if the literature refers to pap or mammography screening.

On page 10 it states that South Dakota reports the fewest percentage (1.8%) of women with an inability to work, but on page 11 it states that Hawaii reports the lowest number (1.9%) of women with an inability to work. Are these the same thing, since they both are given as percentages but a number is claimed for Hawaii? This causes confusion, please clarify.

The last paragraph in the limitations section would be better understood in the methods section where the statistical methods are discussed.

For Tables 1 and 2, are rates what is reported? Rates generally refer to a proportion such as x cases per 1,000 or 100,000. The numbers present are reported as percentages, which are easily understood, but are not the same as rates.

Reviewer #3: This is a nicely written paper on a secondary analysis of cancer screening rates for women with diabetes using a publicly available epidemiological data set. The paper lacks a Methods Section. There are some problems with the methods and analyses (see below). On the whole, these seem correctable.

1. In the Methods Section, BFRSS questions are mentioned before the BFRSS is described. Seems better to describe the BFRSS before the PRECEDE Model.
2. In the Sample Section, subsets of the 2003 BFRSS were used, but the description of the BFRSS does not indicate what years this study was conducted. It should. Also, if 2003 is not the only year, a justification is needed for restricting to 2003. A description of the sampling procedures used in this study is also required since those procedures may have an effect on how the data should be analyzed (see Issue 6).
3. Page 4, lines 4-6, the description of the participants of the current study is not precisely stated. It says first that only women with diabetes were considered, but how then can they be compared to other women? Shouldn't the participants be all women for which the response to the core question is not missing? If pregnant women with diabetes are excluded, why not also exclude non-diabetic pregnant women?
4. Guam seems to be considered one of the 13 "states" using the optional model so why not include results for Guam in the analysis if those data have a reasonable size?
5. Why restrict to those states choosing the optional Women's Health Module? Is this the part of the survey with questions on cancer screening? If not, the restriction seems like discarding useful data.
6. There is no Methods Section. One is needed describing in detail the methods to be used, especially the sampling weights. How are they derived and why it is appropriate to use them with standard chi-square and logistic regression methods rather than with methods which adjust for correlation within sampling clusters.
7. Page 5, lines 5-6. residual analysis is not typically used as part of chi-square analysis. How this was done should be described more precisely in the Methods Section.
8. Page 5, explain in the Methods Section, precisely what is meant by "no multicollinearity was demonstrated".
9. On page 5-6, it states that the "number of days in the past 30 viewed as poor health ..." is excluded from the analyses. This is not true. It was considered on page 8 in analyses of PAP screening rates and then excluded. The sentence on page 6, lines 11-12 is also not correct. While the PRECEDE variables were initially included on the basis of theory, the variable "number of days in the past 30 viewed as poor health ..." was dropped, supposedly because of overlap with other factors of the model but in reality because it was non-significant. If it is left in the model, are any of the reported significances no longer significant? That would justify its exclusion on the basis of masking the effects of other related factors, but that would also be like using the automated selection procedures it says on page 6 are not used in the analysis. If there are no differences in significances with or without this variable in the model, why bother to exclude it?
10. Page 6, line 23, how can an F score of 1.5 fall in a confidence interval of 1.87-41537.70? Are 95% confidence intervals used?
11. Page 11, line 9, the term "outliers" is inappropriately used.
12. Page 16, the last paragraph contains for the first time in the paper information about the handling of missing predictor values for the logistic regression analyses. How missing data was handled should have been presented in the Methods Section. A description of the extent of missing values for the predictor variables is also needed, especially since they seem to have been extensive.

CHECKLIST FOR STYLE

TITLE PAGE --

Supply professional title for each author.

ABSTRACT --

Supply 2-3 keywords; use MeSH guidelines if possible.

REFERENCES --

Update REFERENCE LIST using APA 5th Ed. format. In particular,

Alphabetize the Green reference.

Check the 4th author in Fontana cit/ref (Love or Richard?).

Update IN-TEXT CITATIONS using APA 5th Ed. Format (The Publication Manual of the American Psychological Association, 5th Edition, pp. 208-209). In particular:

For 6 or more authors, use only the first author's name with et al. (e.g., (e.g., Beckman...2001; Friedman...2002).

Check the 4th author in Fontana cit/ref -- (Love or Richard?).

Add in-text citations of the following references to the reference list: Nosek...2001

The following reference is not cited in the text. Remove from reference list or add to text: Stata Corp. (2003).

TABLES -- Format tables according to APA 5th Ed: Remove lines around and/or inside the tables.