Author's response to reviews

Title: Feasibility of Identifying Families for Genetic Studies of Birth Defects Using The National Health Interview Survey

Authors:

Dr Diego F Wyszynski (dfw@bu.edu)
Vikki G Nolan (vnolan@bu.edu)

Version: 4 Date: 2 Apr 2004

Reviewer 1:

Reviewer's report:

Review of CLP manuscript
The authors have revised this manuscript but there are still problems:

1. Study question varies by section of manuscript.

Abstract: "The purpose of this study was to determine whether the NHIS is a useful source to identify informative families ----- defects."
Background: The objective of this paper (should be study) is to identify informative families for genetic studies ----- dataset.
Discussion: The objectives were to identify households by geographical region.
The authors need to be consistent throughout the paper.

We agree with the reviewer and modified the stated objectives so that all three mentions of the study purpose are consistent.

2. The authors estimated the number of birth defects in different categories based on US census. How many would they expect based on the subset of 28,000 households included? This refers to power calculations.

The estimated prevalence rates were calculated based on the 28,000 households that were included in the NHIS sample. These rates per 10,000 households were then applied to the US census to estimate the projected total for the US.

3. Methods: Author confuse percentage and odds ratio in second paragraph.

The specific percentages were added to the paragraph to minimize confusion.

Third paragraph - table 3 single non-familial is highest in South and single familial is highest in Midwest and West.

The third paragraph is referring to table 3 not table 2. In table 3, the Western states do appear to have the highest rates of households with single familial and non-familial birth defects.

4. Discussion: The authors should discuss what their data shows, both the positive and negative aspects. They should discuss why they used it, how it has been used before and what they expected and why they did not find that. Right now it reads like a list. The discussion needs to be rewritten to
give it a better flow.

The discussion section has been revised.

I still suggest shortening this paper to a letter or very short report of the efficiency of using this data set.

Reviewer's report

Title: Feasibility of Identifying Families for Genetic Studies of Birth Defects Using The National Health Interview Survey

General

Analyses using the NHIS-D are very complex and difficult. The authors are to be commended for continuing their efforts to explore the usefulness of the survey for novel applications.

The authors have responded adequately to most of the concerns raised in the first review. The remaining concerns are described below.

This paper does address an important issue (how to do case finding for genetic studies and whether population based surveys offer useful assistance in this effort).

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

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

One remaining concern is that on Tables 2 - 4 the authors estimate projected US totals without taking the NHIS weights into account. For example, on Table 3, the authors use the unweighted rates per 10,000 from the NHIS to project the number of households in the U.S. in each of the four groups. This approach fails to account for the complex multistage probability sample of the NHIS. According to the user's guide for the NHIS-D, (NCHS, 1994), the weights are needed to adjust for the probability of selection (based on PSU, segment and household), household non-response adjustment within segment, and for the first-stage ratio adjustment for race and residence classes, (for analyses at the household level); and for post-stratification age, sex, race adjustments (for analyses at the person level).

Since the authors used Stata to conduct their analysis and the newest versions of Stata do have the capability of correctly accounting for variability in complex samples, they could use the household record weight to properly estimate the number of U.S. households a sampled household represents. This is preferable for two reasons. First, it properly accounts for the complex sampling design of the NHIS. Second, it provides an estimate based on the same year the data were collected instead of 5-6 years later.

I would strongly recommend that for the individual level data presented on Table 2, the authors apply the individual weights and for the household characteristics reported in Tables 2-4, they apply the household weights from the household file. Given the lack of statistical significance for differences between the four groups, I would further recommend that the authors describe the characteristics of all households together and not report odds ratios. If they wish to present odds ratios, the tests of differences must be weighted so that the
variance estimates (including the confidence intervals) are not biased.

An alternative would be to drop Tables 2-4 altogether and to simply report the results without trying to apply them to the U.S. population as a whole. Since the main finding of the report is that the NHIS doesn't work to accomplish the outcome the authors hoped, such an approach would retain the key finding without requiring analyses to be rerun.

The purpose of our study was to describe population trends rather than to calculate precise estimates of prevalence therefore we did not use a weighting strategy. A sentence stating this has been added to the methods section.

My other remaining concern is that each household in the sample responded to only one of the six condition lists. That means that if the household was assigned to condition list 2 they would have been asked about cleft palate or harelip but they would not have been asked about congenital heart disease (since that condition appears on list 6). One of the major purposes of the paper is to identify the number of families that have more than one member with birth defects, but families with members who had two different categories of birth defects would have been under-identified by the selected approach. At best, the methodology reported in the paper would identify families with two or more members who had birth defects in the same category (e.g., skin and musculoskeletal, sensory or impairments, digestive, miscellaneous, circulatory or respiratory). This concern can be responded to by noting the problem in the methodology section and exploring its potential consequence in the discussion section.

A sentence describing this particular limitation was added to the discussion section.

------------------------------------------------------------------------------------------------------------------
Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)
I have no edits of this type to suggest.
------------------------------------------------------------------------------------------------------------------
Discretionary Revisions (which the author can choose to ignore)
I have no edits of this type to suggest.

Reference:

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: Acceptable
Statistical review: No
Declaration of competing interests:
None