Reviewer's report

Title: 118 SNPs of Folate-Related Genes and Risks of Selected Congenital Anomalies

Version: 1 Date: 9 February 2009

Reviewer: James L Mills

Reviewer's report:

Shaw et al.
118 SNPs of Folate-Related Genes and Risks for Selected Congenital Anomalies
Reviewer: James L. Mills, MD, MS

In this paper Shaw and colleagues search for associations between 118 SNPs in folate pathways and neural tube defects (NTDs) or cono-truncal heart defects. They use population based data on births in California.

This is a worthwhile topic for investigation and the researchers provide interesting data. The strengths of the study are the large number of variants examined, the representative sample selected for investigation and the moderately large number of subjects.

The analysis is reasonable with the exception of a few issues that will be discussed below. The conclusions are appropriately cautious. The paper is clearly written. Dr. Shaw's inimitable style is evident in sentences that begin, “Eligible were live born infants…” and “Included were 259 infants…”.

The authors identify almost all the limitations in the Discussion section.

Suggestions:

Note that these fall into the category of Minor Essential Revisions because I believe that the authors will respond appropriately.

1. Most readers will know these variants by the polymorphisms, e.g. MTHFR 677C->T. Those should be used in the text. The rs numbers are helpful, but many people will not be able to relate them to the polymorphisms.
2. In the abstract and in the results section, confidence intervals should be included with the odds ratios.
3. In the abstract conclusions the authors should note that the haplotype findings could be due to chance.
4. My major concern is that the NTD case population is quite different from the control population in race/ethnicity. The authors are aware of this problem. They say that they did not observe evidence that risk patterns were confounded by this difference. These data should be presented. In fact, the regression analysis including race/ethnicity should probably be the major analysis presented in the
paper (results and tables). If the logistic regression results are presented in Table 3, that should be noted in the table. If population stratification can be excluded, it will make the findings much stronger.

5. The authors note briefly in the limitations paragraph that multiple comparisons were made. They should expand this to provide some estimate as to what findings, if any, would be statistically significant if correction for multiple comparisons were performed. This does not have to be done in a way that dismisses all the findings with a p value below 0.05. Correcting by standard methods may be too severe. It does, however, have to indicate that p values less than 0.05 cannot be taken at face value when hundreds of comparisons are made.

6. At the end of the results section, the authors report on haplotype analyses stratified by race. Did the other positive results in the previous paragraph become non-significant in the stratified analyses? This needs to be clarified.

7. I have two comments regarding MTHFD1, one of several genes that we have studied. The first is that we (ref.39) actually performed two studies. The second, referenced in the paper, is the confirmatory study. So it would be helpful to indicate that this gene has been confirmed to be a maternal risk factor in the Irish population. The second is that this study could have found a modest effect in cases when there is a true effect in mothers. That point should be mentioned.

8. On table 4, I was not able to discern why some rows were highlighted.

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

**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.