**Supplementary Materials**

**Probabilistic Bias Analysis for Outcome Misclassification**

We conducted a quantitative probabilistic bias analysis to investigate the potential effects of misclassification of self-reported urinary tract infection (UTI) on study results. We used a free, online tool (<https://sites.google.com/site/biasanalysis/>) that performs probabilistic bias analysis for outcome misclassification using the crude effect estimate. The methods used in this tool are explained in full detail in its accompanying textbook.1

Because we had no data on sensitivity (Se) and specificity (Sp) of UTI classification among NBDPS participants, we had to use an estimate from the published literature. We found a study that compared self-reported UTI in pregnancy to medical records among participants in the Pregnancy Risk Assessment Monitoring System in Vermont and New York.2 Se was 0.48 in Vermont and 0.51 in New York; Sp was 0.89 in Vermont and 0.88 in New York.

When we used these Se and Sp estimates in our bias analyses, we obtained impossible results (negative prevalence ratios [PRs]). Increasing Se to 0.55 and Sp to 0.92 fixed the problem. These results indicate that the Se and Sp from the validation study were inappropriate for our population—women in our study likely had more accurate self-report than women in the validation study.

For the probabilistic bias analysis, we created trapezoidal distributions for Se and Sp. Trapezoidal distributions are defined by their minimum, lower mode, upper mode, and maximum. For Se, we used a trapezoidal distribution of (0.55, 0.60, 0.65, 0.70), based on a minimum Se of 0.55 and assigning a distance between the minimum, modes, and maximum of 0.05 each. Our choice of 0.05 was somewhat arbitrary; we had no information to guide us and so we chose a range of Se values that seemed plausible. For Sp, we used a trapezoidal distribution of (0.92, 0.94, 0.96, 0.98) based on the minimum Sp of 0.92 and a distance of 0.02 between the values. Our choice of 0.02 was also arbitrary but was limited by our need to keep Sp at or below 1.

We conducted three probabilistic analyses. In the first, exposed and unexposed women were assigned the same trapezoidal distributions for Se and Sp to simulate nondifferential misclassification. In the second analysis, we shifted the trapezoidal distributions for Se and Sp upward by 0.10 and 0.02 for the exposed women, to (0.65, 0.70, 0.75, 0.80) and (0.94, 0.96, 0.98, 1.00). This shift simulated differential misclassification in which exposed women had more accurate self-report than unexposed women. In the third analysis, we shifted the trapezoidal distributions for Se and Sp upward by 0.10 and 0.02 for the unexposed women to simulate unexposed women having more accurate self-report. Our choice of 0.10 for the upward shift for Se was arbitrary given our lack of information about the accuracy of UTI self-report in pregnancy, but was chosen because it represented a plausible range of Se values. 0.02 was chosen for Sp so that its value would not exceed 1.

To adjust for outcome misclassification, the tool sampled Se and Sp values from the resultant distributions and used them to back-calculate the PR that would have been observed in the absence of outcome misclassification. Random error was added by sampling an error term from a normal distribution. For both Se and Sp, the values were chosen for exposed and unexposed women with 0.80 correlation because all women answered the same questionnaire and should therefore have similarities in reporting. The sampling and back-calculation were repeated 1,000 times to generate a distribution of plausible PRs. We reported results as the median PR from the distribution and the 95% simulation interval—an interval that contains 95% of the PRs from simulation and that is defined by the 2.5th and 97.5th percentiles of the PR distribution.

**REFERENCES**

1. Lash TL, Fox MP, Fink AK. Applying quantitative bias analysis to epidemiologic data. 2009. New York: Springer.
2. Dietz P, Bombard J, Mulready-Ward C, Gauthier J, Sackoff J, Brozicevic P, Gambatese M, Nyland-Funke M, England L, Harrison L, Taylor A. Validation of self-reported maternal and infant health indicators in the Pregnancy Risk Assessment Monitoring System. Matern Child Health J 2014; 18(10):2489-2498.