Subject: Re: Missing data Posted by Bridgette-DHS on Wed, 05 Mar 2014 17:09:47 GMT View Forum Message <> Reply to Message

Following is another response from one of our DHS experts, Shea Rutstein.

In the DHS we do not impute whether each child received a vaccination. If a vaccine dose is not recorded on the the child's immunization card, then the mother is asked whether the child received the vaccination. For DPT, the questions asked are in the attached file.

The age at which the vaccine is given is assumed to be the same as that for which dates have been given (an aggregate assignment, not individual, done during tabulation). Since the outcome of whether a child is given a vaccination is dichotomous, either logistic or probit regression is appropriate. The predicted value will be the probability that the child was given the vaccination. I would combine the several imputations to get the average probability for each child and then randomly select a number so that you assign 1 or 0 (got vaccination or not) according to the number selected, e.g. if the probability of receiving DPT1 is 0.60 then a randomly selected number between 0 and 5 would indicate that the vaccine was given and between 6 and 9 would indicate that the vaccine was given and between 6 and 9 would indicate that the vaccine was given and between 6 and 9 would indicate that the vaccine was given and between 6 and 9 would indicate that the vaccine was given and between 6 and 9 would indicate that the vaccine was given and between 6 and 9 would indicate that the vaccine was given and between 6 and 9 would indicate that the vaccine was given and between 6 and 9 would indicate that the vaccine was given and between 6 and 9 would indicate that the vaccine was given and between 6 and 9 would indicate that the vaccine was given and between 6 and 9 would indicate that the vaccine was given and between 6 and 9 would indicate that the vaccine was given and between 6 and 9 would indicate that the vaccine was given and between 6 and 9 would indicate that the vaccine was given and between 6 and 9 would indicate that the vaccine was given and between 6 and 9 would indicate that the vaccine was given and between 6 and 9 would indicate that the vaccine was given and between 6 and 9 would indicate that the vaccine was given and between 6 and 9 would indicate that the vaccine was given and between 6 and 9 would indicate that the vaccine was given and between 6 and 9 would indicate that the vaccine was given and between 6

Another way to go about it would be to use hot deck imputation according to the characteristics correlated with each vaccination, such as child's sex, birth order, age, area of residence, province, place of birth, wealth quintile, etc. but the list of variables may be long.

As I understand it, Rubin's procedure for combining multiple imputations is to produce a more robust standard error by taking account of both interval variance and variance between the estimates. In this case, I would take the calculated probability of each child from the estimating equation and then randomly vary by selecting a deviation using the normal distribution of the standard error, and then apply the adjusted probability for each child to a randomly selected number to determine wether the vaccine was given.

Let me know if this helps.

Shea

File Attachments
1) missing data.jpg, downloaded 1014 times