

A preliminary study of active compared with passive imputation of missing body mass index values among non-Hispanic white youths^{1–4}

David A Wagstaff, Sibylle Kranz, and Ofer Harel

ABSTRACT

Background: Addressing missing data on body weight, height, or both is a challenge many researchers face. In calculating the body mass index (BMI) of study participants, researchers need to impute the missing data.

Objective: A multiple imputation through a chained equations approach was used to determine whether one should first impute the missing anthropometric data and then calculate BMI or use an imputation model to obtain BMI.

Design: The present study used computer simulation to address the question of how to calculate BMI when there is missing data on weight and height. The simulated data reflected data gathered on non-Hispanic white youths ($n = 905$) aged 2–18 y, who participated in the 1999–2000 National Health and Nutrition Examination Survey (NHANES).

Results: The simulation indicated that it made little difference in the accuracy with which the youths' mean BMIs were estimated when the data were missing completely at random. However, the use of a model to impute BMI was favored slightly when the data were missing at random and the imputation model included the variable used to determine missingness.

Conclusion: The present findings extend the use of passive imputation and the use of multiple imputation through a chained equations approach to an area of critical public health importance. *Am J Clin Nutr* 2009;89:1025–30.

INTRODUCTION

Overweight and obesity have reached epidemic proportions in many countries (1–4). Of the measures used to assess an individual's body fat, the most prominent is body mass index (BMI), which is calculated as weight (in kg)/height² (in m) (1, 5, 6). The latter cannot be calculated when either measurement is missing, and failure to address the missing measurements may bias the estimation. To minimize bias, many researchers have used multiple imputation (7, 8).

Multiple imputation has been widely used since Rubin proposed his framework (9–13). Many researchers have modified these methods; therefore, frequent reviews have become necessary (14–19). Rubin's framework identifies 3 missing data mechanisms (13). When data are missing completely at random (MCAR), the probability of nonresponse cannot be predicted by any variable, observed or unobserved, and participants with incomplete data constitute a random sample of all participants.

When data are missing at random (MAR), the probability of nonresponse depends on observed and missing variables. Finally, when data are missing not at random, the probability of nonresponse depends on observed variables.

To date, the literature on multiple imputation has not addressed a number of basic questions (20). One such question is, "Should researchers impute the components needed to calculate an index or should they impute the index?" Two exceptions include studies conducted by Liu et al (21) and Gmel (22). Liu et al had missing data on one or more 12-item Short Form Health Survey (SF-12) items and wanted to calculate the 2 SF-12 component scores. They used computer simulation to compare the performance of a "simple" imputation model that used the available SF-12 item scores with that of an "enhanced" model that supplemented the available scores with other participant data. The enhanced model performed better when ≥ 6 SF-12 items were missing; the simple model was adequate when the number of missing items was < 6 . Because Gmel's respondents had missing data on ≥ 1 of 8 components, Gmel could not calculate a summary measure of alcohol consumption. Gmel imputed item scores because the distribution displayed by his items was most consistent in his multivariate normal imputation model.

If data were missing on a few items, most researchers imputed the missing component scores. This practice would be consistent with that of Liu et al (21). When 2 variables (eg, weight and height) are used to calculate an index such as BMI, will it matter whether the researcher imputes the missing data or simply imputes the index? The present study used computer simulation to address this question using simulated data consistent with that

¹ From the Department of Human Development and Family Studies, College of Health and Human Development, The Pennsylvania State University, University Park, PA (DAW); the Department of Statistics, University of Connecticut, Storrs, CT (OH); and the Department of Nutrition and Dietetics, East Carolina University, Greenville, NC (SK).

² The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Child Health and Human Development or the National Institutes of Health.

³ Supported in part by grant R03HD050239 from the National Institute of Child Health and Human Development.

⁴ Reprints not available. Address correspondence to DA Wagstaff, S153 Henderson Building, Department of Human Development and Family Studies, College of Health and Human Development, The Pennsylvania State University, University Park, PA 16802. E-mail: daw22@psu.edu.

Received September 19, 2008. Accepted for publication January 30, 2009.

First published online February 25, 2009; doi: 10.3945/ajcn.2008.26995.

reported for non-Hispanic white youths participating in a national survey. It should be noted that these data were used for illustration purposes only.

SUBJECTS AND METHODS

The National Center of Health Statistics, a unit of the Centers for Disease Control and Prevention, periodically collects nationally representative data on the health and nutrition status of the US population by using the National Health and Nutrition Examination Survey (NHANES) (23). Trained personnel interview family members and collect data on the participants' sociodemographic characteristics, diet, and health. A proxy is interviewed if the individual is aged ≤ 15 y or cannot provide a report. Participating individuals are invited to visit a mobile center for standardized medical examinations that include anthropometric measurements. The present study is based on data obtained from 905 non-Hispanic white youths between the ages of 2 and 18 y when they participated in the 1999–2000 NHANES.

Measures

Age (y) was assessed by asking the individual or the proxy for the sampled person's exact date of birth. Standing height (cm) was measured to the nearest tenth of a centimeter with a fixed stadiometer with a vertical backboard and a moveable headboard. Abdominal waist circumference (cm) was measured to the nearest tenth of a centimeter with a steel measuring tape according to standard procedures. Weight (kg) was measured as the individual stood on a digital according to standard procedures. Two scales were used if the individual weighed >440 lb (198 kg). The individual's weight was set equal to missing if the individual had undergone a limb amputation.

Poverty-income ratio (PIR) is the ratio of self-reported total household income to the number of adults and children in the family. PIR values <1.00 are below the official poverty threshold; PIR values >5 were set equal to 5. BMI (in kg/m^2) is a relative measure of body fat that is calculated by dividing weight by height squared. For children and adolescents, classification as normal/healthy weight, overweight, or obese is based on sex- and age-specific BMI-for-age growth charts prepared by the Centers for Disease Control and Prevention.

Procedures

The present study mirrored the steps that Ezzati-Rice et al (24) used to compare different imputation techniques with NHANES 1999–2000 data. However, it differed from Ezzati-Rice et al by restricting attention to 6 variables: age, height, weight, PIR, waist circumference, and BMI. Summary statistics for these variables and the percentage of missing data on weight and height are shown in **Table 1**. All computations were done using Stata software (version 9), a general purpose statistical package (25). To initiate the simulation, Stata's hotdeck program was used to identify plausible target values for the 6 means and the corresponding variance-covariance matrix. These summary statistics were used by Stata's drawnorm program to generate multivariate normal samples. The latter were generated because of the widespread availability of multivariate normal generators. Once target values were identified, the following 4 steps were executed:

- Step 1: Draw a multivariate normal sample of 1080 observations from the hypothetical population, delete all cases for which generated values are ≤ 0 , randomly sort the remaining records and delete all records in excess of $n = 905$, and proceed to step 2 if $n = 905$.
- Step 2: Impose the missing height-weight data pattern that was observed among the 905 NHANES participants by using an MCAR (MAR) missing data mechanism.
- Step 3: Use Stata's ICE program to obtain 10 multiply imputed data sets (26–28). This imputation method, which is described later in greater detail, uses a sequence of regression models to impute missing values conditional on other predictors.
- Step 4: Use Stata's MIM program (29) to apply Rubin's rules (13) and obtain the mean, SE, and df for the combined estimates of height, weight, and BMI, and then write the 9 values to an output file for further analyses.

Steps 1–4 were repeated 5000 times with an MCAR missing data mechanism and with an MAR missing data mechanism. To implement an MCAR missing data mechanism, 905 records were generated, randomly sorted, and merged with a file that consisted of two 0/1 variables. If the value was 1, the corresponding value of weight or height was set equal to missing; if either weight or height was set equal to missing, the corresponding BMI value was set equal to missing. To implement an MAR missing data mechanism, the 905 rows of the 0/1 indicators were shifted in the secondary file until the probability of missing height was predicted by age, which was fully observed. (In the 1999–2000 NHANES, age was observed for all 905 participating non-Hispanic white youths.)

TABLE 1
Selected demographic characteristics and anthropometric measurements of non-Hispanic white youths participating in the 1999–2000 National Health and Nutrition Examination Survey ($n = 905$)

Variable	Name	No. of cases	Mean	SD	Percentage	Minimum	Maximum
Age (y)	ridageyr	905	10.23	5.11	—	2	18
Poverty-income ratio	indfmpir	817	2.83	1.62	—	0	5
Weight (kg)	bmxtwt	841	42.89	23.80	—	10.90	142.20
Height (m)	bmxtht	835	141.62	28.62	—	83.70	193.90
BMI (kg/m^2)	bmxbmi	834	19.55	4.80	—	12.15	46.66
Waist circumference (cm)	bmxtwaist	800	68.09	15.97	—	42.50	142.30
BMI status	—	—	—	—	—	—	—
Missing weight only	—	—	—	—	0.12	—	—
Missing height only	—	—	—	—	0.78	—	—
Missing weight and height	—	—	—	—	6.95	—	—
Complete weight and height	—	—	—	—	92.15	—	—

To impute BMI passively, the following 2 prediction equations were used to actively impute values for height and weight:

$$\begin{aligned} \text{Weight} = & \beta_0 + \beta_1 \text{ age} + \beta_2 \text{ height} \\ & + \beta_3 \text{ waist} + \beta_4 \text{ PIR} + \text{residual} \end{aligned} \quad (1)$$

$$\text{Height} = \beta_0 + \beta_1 \text{ weight} + \beta_2 \text{ age} + \beta_3 \text{ waist} + \text{residual} \quad (2)$$

Then, after having obtained imputed values for weight and height, an equation preceded with the keyword *passive* was used to define the relation between BMI, weight, and height:

$$\text{BMI} = \text{weight}/[(\text{height}/100) \times (\text{height}/100)] \quad (3)$$

To actively impute BMI, the following prediction equation was used:

$$\text{BMI} = \beta_0 + \beta_1 \text{ age} + \beta_2 \text{ waist} + \beta_3 \text{ PIR} + \text{residual} \quad (4)$$

Passive imputation refers to methods that were built into the multiple imputation through a chained equations (MICE) approach by van Buuren and Oudshoorn (30) to ensure that transformed variables (BMI in the present study) are always in sync with the variables (weight and height) that are used to define them. Active imputation refers to the usual methods and models used to impute data with this approach.

MICE combines the most attractive features of multivariate normal imputation and the more flexible regression-based approaches to imputation. The Stata program that implements this approach is called ICE (26–28). With this approach, the user specifies a conditional distribution for the missing data on each variable. For example, using Equations 1 and 2, linear regression models were specified for weight and for height. Stata's ICE program cycles through the system of equations until all variables have complete data. (Cycles handle incomplete data on any specified predictor, and users can use a dry run option to see the models that are fitted.) The critical assumption on which this approach is built is that a multivariate distribution exists from which the specified or implied conditional distributions can be derived and that one can use iterative Gibbs sampling to draw random values from these conditional distributions. Both Raghunathan et al's (31) sequential regression approach and a multiple imputation through a chained equations approach cycle through a sequence of regression models to obtain imputed values. In comparative

studies of various imputation approaches, the MICE approach has performed well (32, 33).

RESULTS

Summary statistics for the sample of non-Hispanic white youths ($n = 905$) who participated in the 1999–2000 NHANES are shown in Table 1. According to the US growth chart, 69.4% of the youths were of normal weight, 15.3% were overweight, 7.5% were obese, and 7.9% could not be classified because of missing weight and/or height measurements.

The results from the 3 test runs are summarized in the first 3 rows of **Table 2**. The latter were used to verify the performance of the multivariate normal data generator and the code used to impose the missing data mechanisms and create the incomplete samples. No data were imputed. Each test run consisted of 10 trials. For each trial, 5000 samples ($n = 1080$) were generated. Because negative values of age, PIR, weight, height, and waist circumference are not plausible, any records with negative values that were generated at step 1 were dropped before proceeding to step 2. Test 1 was used to estimate the maximum number of records with one or more negative values that would be generated during a trial (and, hence, the number of records that had to be generated for each sample to ensure that each sample consisted of 905 records with plausible values). Across the 10 trials of test 1, the average number of records with negative values ranged from 43 to 110.

The 3 test runs were also used to estimate target values for mean BMI when no missing data mechanism was imposed, when an MCAR missing data mechanism was imposed, and when an MAR missing data mechanism was imposed. The absolute and relative errors for the various runs are shown in the upper portion of **Table 3**. Given the target value of 20.53, the absolute error was calculated as $100 \times (\hat{\theta} - \theta)\%$, and the relative error was calculated as $100 \times [(\hat{\theta} - \theta)/\theta]\%$. A fourth test (data not shown) was conducted to verify that control had been achieved over the relation between the BMI missing data indicator and age under the MAR missing data mechanism. The coefficient for the regression of the missing BMI indicator on the values of age generated under the MAR missing data mechanism only varied by 0.01% over 10 trials.

To assess the extent to which the MCAR findings were dependent on the use of Stata's multivariate normal random number generator and the subsequent exclusion of negative values, 2 additional test runs were conducted (data not shown). Each run consisted of one trial for which 5000 missing data samples were

TABLE 2
Summary statistics for the test and simulation runs¹

Unit	No. of values	Mean	SD	Minimum	Maximum	Remarks
Test 1	50,000	20.53	0.21	19.58	21.41	MVN DGP: no negative values; no missing data mechanism
Test 2	50,000	20.53	0.22	19.64	21.36	MVN DGP: no negative values; MCAR: no imputation
Test 3	50,000	20.51	0.22	19.59	21.40	MVN DGP: no negative values; MAR: no imputation
Simulation 1	50,000	20.52	0.21	19.77	21.34	MVN DGP: no negative values; MCAR: active imputation of weight and height and passive imputation of BMI
Simulation 2	50,000	20.39	0.22	19.61	21.18	MVN DGP: no negative values; MAR: active imputation of weight and height and passive imputation of BMI
Simulation 3	50,000	20.53	0.21	19.72	21.21	MVN DGP: no negative values; MCAR: active imputation of BMI
Simulation 4	50,000	20.56	0.22	19.73	21.65	MVN DGP: no negative values; MAR: active imputation of BMI

¹ MVN DGP, multivariate normal data-generating process; MCAR, missing completely at random; MAR, missing at random.

TABLE 3

Absolute and relative error estimating the target BMI value (20.53)

Unit	Combined mean	Absolute error	Relative error	Remarks
Test 1	20.53	0.0	0.0	No imputation
Test 2	20.53	0.0	0.0	MCAR MDM; no imputation
Test 3	20.51	-2.0	-0.10	MAR MDM; no imputation
Simulation 1	20.52	-1.0	-0.05	MCAR MDM; passive imputation
Simulation 2	20.39	-14.0	-0.68	MAR MDM; passive imputation
Simulation 3	20.53	0.0	0.0	MCAR MDM; active imputation
Simulation 4	20.56	3.0	0.15	MAR MDM; active imputation

¹MCAR, missing completely at random; MAR, missing at random; MDM, missing data mechanism.

created and processed. Each run started with the same hotdeck-completed data set ($n = 905$), for which the mean BMI was 19.4692. On each run, the records were randomly sorted, and the weight and/or height measurements were set equal to missing after merging with a randomly sorted file with two 0/1 indicators for the NHANES missing weight-height data pattern. Stata's ICE program was used to return 10 multiply imputed data sets, and MIM was used to process these data sets and to obtain a combined mean and SE for the BMI scores that had been imputed passively (actively). When BMI scores were imputed passively (Equations 1–3), the absolute error was 0.98% and the relative error was 0.05%; when BMI scores were imputed actively (Equation 4), the absolute error was -0.04% and the relative error was 0.0%. These values were comparable with those reported when the data were generated with a multivariate normal model and records with negative values were excluded.

On completion of the test runs, the present study moved to the simulation phase. With each of the 4 simulation runs, the total sample size ($n = 905$) on entry to step 2 and the missing weight-height data patterns on exit from step 2 were identical. Moreover, the missing weight-height data patterns were the same as the patterns observed in the NHANES sample (Table 1). Finally, 5000 samples were generated for each simulation run, a missing data process was imposed on each sample, and 10 multiply imputed data sets were created and analyzed.

With simulation 1, an MCAR missing data mechanism was imposed after each sample was generated. Then, ICE used Equations 1 and 2 to actively impute plausible values for weight and height; ICE used Equation 3 to passively impute missing BMI scores. With each call to ICE, 10 multiply imputed data sets were created. After the imputed data sets were written to disk, another Stata program, MIM (28), was used to process the multiply imputed data sets and to obtain combined estimates and SEs for weight, height, and BMI.

With simulation 2, an MAR missing data mechanism was imposed. Stata's ICE program was used to actively impute plausible values for weight and height and to passively impute plausible values for BMI. Ten multiply imputed data sets were created with each call to ICE, and MIM was used to obtain combined estimates and SEs for weight, height, and BMI.

With simulation 3, an MCAR missing data mechanism was imposed. Then, ICE used Equation 4 to actively impute BMI and return 10 multiply imputed data sets. After the data sets were written to disk, MIM was used to obtain combined estimates and SEs for BMI.

With simulation 4, an MAR missing data mechanism was imposed, and ICE was called to actively impute BMI and return

10 multiply imputed data sets. After the data sets were written to disk, MIM was used to obtain combined estimates and SEs for BMI. The absolute and relative errors observed for the 4 simulation runs are shown in the lower portion of Table 3.

DISCUSSION

Many data sets have missing or implausible data for weight and height because of nonresponse, measurement error, or data entry problems. The results of the present study indicate that researchers can use passive or active imputation to address missing data on height and/or weight and obtain individuals' BMI values when the data are MCAR or MAR. Precise estimates of the mean BMI were obtained with passive and with active imputation when the missing data mechanism was MCAR. Less, but still reasonably, precise estimates of BMI were obtained with passive and with active imputation when the missing data mechanism was MAR. However, it must be stated that the imputation model included age—the variable used to impose the MAR missing data process. An obvious advantage of a missing data simulation is that researchers know and control the factors that determine missingness. Meeting this condition is far more challenging when researchers collect data from study participants.

Limitations

As with all simulation-based research on missing data, the present findings are limited to the number and variable types and to the missing data percentages and patterns that were considered. By no means did the present study attempt to be comprehensive or mimic any frequently occurring missing data situation. The authors restricted their attention to anthropometric data reported for youths participating in the 1999–2000 NHANES because they were working with the NHANES data and they needed BMI scores that were not observed for all youths. The authors only considered a limited number of variables in the present study because this limited number provided them with the means to assess their question, “Should one impute the missing data on weight and/or height and then calculate BMI or should one use an imputation model to obtain BMI?”

Conclusions

Rubin (10–14) proposed a framework for estimating population parameters with incomplete data. More importantly, he and subsequent researchers developed software that can be used to create and analyze multiply imputed data sets (34). As these tools have matured, researchers have sought to address practical

“how-to” questions. The present study used computer simulation to ask, Should one calculate BMI from imputed values of weight and/or height, or should one use an imputation model to impute BMI? The answer to this particular question is important for ≥ 3 reasons. First, BMI is associated with elevated morbidity and mortality. Additionally, BMI is used to mark a country’s progress toward ameliorating the current epidemic of overweight and obesity. Second, despite the relative ease and accuracy with which an individual’s weight and height can be measured by a trained individual, missing BMI data are found in well-executed studies. Third, variables such as waist circumference can be collected by an individual who has been trained to use a tape measure, are highly correlated with individuals’ BMIs, and can function effectively in an imputation model for BMI.

Did it matter whether BMI was imputed passively or actively? The results of the present study show that it made very little difference whether BMI was imputed passively or actively when individuals’ weight and/or height measurements were MCAR. On the other hand, active imputation of BMI was favored slightly when the data were MAR. This latter finding is undoubtedly due in part to the fact that the imputation model for BMI included age—the variable that had been used to impose the MAR missing data mechanism.

The finding that multiple imputation’s combined estimator was unbiased when the missing data mechanism was MCAR and when the missing BMI values were imputed passively or actively is consistent with missing data theory and with the findings reported by other researchers. Indeed, estimates obtained by ad hoc methods, such as mean substitution, are unbiased when the data are MCAR. Our second major finding that the combined estimator of mean BMI was biased slightly when the missing data mechanism was MAR and BMI was imputed passively is also consistent with missing data theory and the point that multiple imputation is only as good as the imputation models for weight and height: the dropoff in precision was due to the performance of the imputation models for weight and height. From the literature, it was relatively easy to find an anthropometric measure—waist circumference—that was strongly correlated with BMI and obtained when the youth received a physical examination. Such was not the case for individuals’ weights and heights.

Implications

The present findings generalize to ratios similar to the BMI, such as weight/height, height/ $\sqrt{\text{weight}}$, weight/height³, the waist-hip ratio, and waist circumference-to-height. However, our findings clearly suggest similar simple ratios (ie, ratios of the form u^a/v^b) for which the researcher has collected data that can be used to obtain plausible values. The present findings also extend the use of passive imputation to ratios. Previous examples have included product interaction terms and linear composites.

As substantive researchers confront the challenges posed by missing anthropometric data, additional data collection of measures that would afford imputation of missing data should be considered. The present study indicates that missing BMI values can be imputed passively or actively when the missing data mechanism is ignorable (MCAR or MAR) and when data on a related measure (eg, waist circumference) have been collected.

The authors’ responsibilities were as follows—SK and DAW: conceived the study; DAW: reviewed the literature, conducted the simulation, and con-

tributed to the writing; OH: guided the data simulation methodology and contributed to the writing; and SK: led the project and contributed to the writing. All authors approved this manuscript. None of the authors reported a conflict of interest.

REFERENCES

1. Barlow SE, Expert Committee. Expert committee recommendations regarding the prevention, assessment, and treatment of child and adolescent overweight and obesity: summary report. *Pediatrics* 2007;120:S164–92.
2. Canoy D, Buchan I. Challenges in obesity epidemiology. *Obes Rev* 2007;8(suppl 1):1–11.
3. Reilly JJ. Childhood obesity: an overview. *Child Soc* 2007;21:390–6.
4. Stein CJ, Colditz GA. The epidemic of obesity. *J Clin Endocrinol Metab* 2004;89:2522–5.
5. Al-Lawati JA, Jousilahti P. Body mass index, waist circumference and waist-to-hip ratio cut-off points for categorization of obesity among Omani Arabs. *Public Health Nutr* 2008;11:102–8.
6. Vazquez G, Duval S, Jacobs DR, Silventoinen K. Comparison of body mass index, waist circumference, and waist/hip ratio in predicting incident diabetes: a meta-analysis. *Epidemiol Rev* 2007;29:115–28.
7. Mishra GD, Dobson A. Multiple imputation for body mass index: lessons from the Australian Longitudinal Study on Women’s Health. *Stat Med* 2004;23:3077–87.
8. Nandram B, Han G, Won Choi J. A hierarchical bayesian nonignorable nonresponse model for multinomial data from small areas. *Surv Methodol* 2002;28:145–56.
9. Rubin DB. Inference and missing data. *Biometrika* 1976;63:581–92.
10. Rubin DB. An overview of multiple imputation. In: *Proceedings of the Survey Research Methods Section of the American Statistical Association*. 1988:79–84.
11. Rubin DB. Multiple imputation for nonresponse in surveys. 1st ed. New York, NY: John Wiley and Sons, Inc, 1987.
12. Little RJA. Inference about means for incomplete multivariate data. *Biometrika* 1976;63:593–604.
13. Little RJA, Rubin DB. *Statistical analysis with missing data*. 2nd ed. New York, NY: John Wiley and Sons, Inc, 2002.
14. Rubin DB. Multiple imputation after 18+ years (with discussion). *J Am Stat Assoc* 1996;91:473–89.
15. Schafer JL. Multiple imputation: a primer. *Stat Methods Med Res* 1999;8:3–15.
16. Schafer JL, Graham JW. Missing data: our view of the state of the art. *Psychol Methods* 2002;7:147–77.
17. Donders ART, van der Heijden GJMG, Stijnen T, Moons KGM. Review: a gentle introduction to imputation of missing values. *J Clin Epidemiol* 2006;59:1087–91.
18. Harel O, Zhou X-H. Multiple imputation: review of theory, implementation and software. *Stat Med* 2007;26:3057–77.
19. Kenward MG, Carpenter J. Multiple imputation: current perspectives. *Stat Methods Med Res* 2007;16:199–218.
20. White IR, Wood A, Royston P. Multiple imputation in practice. *Stat Methods Med Res* 2007;16:195–7.
21. Liu H, Hays RD, Adams JL, et al. Imputation of SF-12 health scores for respondents with partially missing data. *Health Serv Res* 2005;40:905–21.
22. Gmel G. Imputation of missing values in the case of a multiple item instrument measuring alcohol consumption. *Stat Med* 2001;20:2369–81.
23. Kuczmarski RJ, Ogden CL, Grummer-Strawn LM, et al. *CDC growth charts: United States*. Hyattsville, MD: National Center for Health Statistics, 2000. (Advance data from vital and health statistics; no. 314.).
24. Ezzati-Rice TM, Khare M, Rubin DB, Little RJA, Schafer JL. A comparison of imputation techniques in the Third National Health and Nutrition Examination Survey. In: *Proceedings of the Survey Research Methods Section of the American Statistical Association*, 1993:303–8. Available from: http://www.amstat.org/sections/srms/proceedings/papers/1993_047.pdf (cited 17 Feb 2009).
25. StataCorp. *Stata statistical software: release 9*. College Station, TX: StataCorp LP, 2005.
26. Royston P. Multiple imputation of missing values. *Stata J* 2004;4:227–41.
27. Royston P. Multiple imputation of missing values: update of ice. *Stata J* 2005;5:527–36.

28. Royston P. Multiple imputation of missing values: further update of ICE, with an emphasis on interval censoring. *Stata J* 2007;7: 445–64.
29. Carlin JB, Galati JC, Royston P. A new framework for managing and analyzing multiply imputed data in Stata. *Stata J* 2008;8:49–67.
30. van Buuren S, Oudshoorn K. Flexible multivariate imputation by MICE. Leiden, Netherlands. Organization for Applied Scientific Research (TNO) 1999; TNO-PG publication number PG/VGZ/99.054. Available from: <http://web.inter.nl.net/users/S.vanBuuren/mi/hmtl/mice.htm> (cited 9 August 2008).
31. Raghunathan TE, Lepkowski JM, van Hoewyk M, Solenberger PW. A multivariate technique for multiply imputing missing values using a sequence of regression models. *Surv Methodol* 2001;27:85–95.
32. Ambler G, Omar RZ, Royston P. A comparison of imputation techniques for handling missing predictor values in a risk model. *Stat Methods Med Res* 2007;16:277–98.
33. Van Buuren S. Multiple imputation of discrete and continuous data by fully conditional specification. *Stat Methods Med Res* 2007;16:219–42.
34. Horton NJ, Kleinman KP. Much ado about nothing. *Am Stat* 2007;61: 79–90.