Title

Increasing the Validity and Efficiency of Blood Pressure Estimates Using Ambulatory and Clinic Measurements and Modern Missing Data Methods

Journal

American Journal of Hypertension, 25(7)

ISSN

0895-7061

Authors

Zawadzki, Matthew J
Graham, John W
Gerin, William

Publication Date

2012-07-01

DOI

10.1038/ajh.2012.40

Peer reviewed

nature publishing group

Increasing the Validity and Efficiency of Blood Pressure Estimates Using Ambulatory and Clinic Measurements and Modern Missing Data Methods

Matthew J. Zawadzki[1], John W. Graham[2] and William Gerin[2]

BACKGROUND
Ambulatory blood pressure monitoring (ABPM) is considered the gold standard for BP measurement, compared to clinic BP measurements (CBP), which are a less valid predictor of target organ damage and cardiovascular events. However, ABPM is considerably more expensive than CBP, leaving BP researchers with a difficult dilemma: Use the less efficient CBP measure, or bear the cost of the more expensive ABPM. Recent developments in missing data methods, notably the two-method measurement (TMM) design, address this problem. With the TMM design, all research participants receive the less expensive CBP measure, but only a random subset receives the more expensive ABPM. The total number of participants must be increased, with additional participants receiving only CBP measurements. Even so, the TMM still reduces costs.

METHODS
We applied the TMM approach, which makes use of a "bias correction" structural equation model, to an empirical data set in which data were available for ABPM and CBP, as well as an echocardiographic measure of left ventricular mass (LVM).

RESULTS
Based on an estimated fivefold difference in cost for using ABPM compared to CBP, we found that statistical power can be considerably increased, or that BP measurement costs can be considerably reduced, when using this planned missing data design.

CONCLUSIONS
These benefits were observed with no loss of predictive validity (i.e., the observed association between BP and LVM). This suggests that the TMM design is a promising technique that in some studies may be able to decrease costs and/or increase one's power to detect effects.

Keywords: *ambulatory blood pressure; blood pressure; clinic blood pressure ; hypertension; missing data; two-method measurement design*

American Journal of Hypertension, advance online publication 19 April 2012; doi:10.1038/ajh.2012.40

The problem of blood pressure measurement

Two or more approaches are often taken to measure blood pressure (BP) as an outcome in empirical research. Measurements taken by the physician, nurse, or technician (clinic BP (CBP)), or taken using ambulatory BP monitoring (ABPM). Although the measures ostensibly tap the same construct (BP), ABPM is considered the "gold standard" for BP measurement, as it is superior to CBP in terms of reliability and construct validity (i.e., ability to predict outcomes). This situation applies not only to BP, but to any construct that is assessed using more than one measure.

CBP measurements. A strength of CBP is that it can be done quickly and inexpensively. Based on clinic readings, clinicians make immediate decisions about a treatment plan, or the need for additional testing. However, these measurements provide only a "snapshot" of the BP when measured in the physician's office or the clinic, circumstances that are atypical of the patient's normal environment. As a result, these measures are subject to biases, notably white coat hypertension, which requires ABPM to detect.

ABPM strengths. ABPM involves the use of a portable device (typically worn for 24 h). Multiple measurements are taken throughout the day and often throughout the night. ABPM has been shown to be a better predictor of target organ damage and cardiovascular events than CBP.[1] In addition, ABPM is the most reliable way to detect masked hypertension, a condition in which patients' BPs are normal in the physician's office, but in the hypertensive range in their normal lives; as a result, they are undiagnosed and untreated.[2] The superior predictive power of ABPM is likely due not only to the increased number of readings it provides, which increases reliability of the measure, but also to its ability to capture the impact of stressors and other environmental factors that occur in daily life and which are likely to affect the BP.[3,4]

ABPM weaknesses. All things being equal, one would use ABPM in preference to CBP. However, although clearly a useful

[1]Department of Psychology, The Pennsylvania State University, University Park, Pennsylvania, USA; [2]Department of Biobehavioral Health, The Pennsylvania State University, University Park, Pennsylvania, USA. Correspondence: William Gerin (wxg17@psu.edu)

Received 9 September 2011; first decision 11 October 2011; accepted 29 February 2012.

tool for research, a drawback to the use of ABPM is its expense. Clinic measurements tend to take no more than 10 min, require little preparation by patient or physician, and require inexpensive equipment. We estimate that ABPM has roughly a fivefold increase in cost compared to CBP (see **Table 1**). These extra costs are found in (i) the time it takes to use the device, including (a) recruiting and instructing patients (e.g., cannot shower or participate in strenuous exercise), (b) instrumenting and validating the device, (c) collecting equipment from participants and downloading data, once the monitoring period has ended, and (d) cleaning and inspecting the device; (ii) the increased cost of the device itself; and (iii) the increased reimbursement needed for research patients.

An undesirable choice. The apparent value of CBP is that it requires far less training, preparation, and expense than ABPM, and one can thus afford to collect data on a larger sample. However, the study results suffer due to the poor construct validity. Thus, researchers must choose between making use of CBP measures knowing their limitations, or bearing the greater expense of ABPM, to obtain desirable levels of statistical power. We offer a third alternative that yields equivalent power, but at decreased expense, to ABPM.

The two-method measurement approach

Two-method measurement (TMM) is a planned missingness design built around the strategy of combining data from a less expensive measure–one with questionable construct validity–and data from a more costly measure (with superior construct validity).[5] At the heart of the TMM approach is the statistical model (**Figure 1**) that we call a bias control model. This model, which is estimated using structural equation modeling methods, partitions the variance of the less expensive measure between a latent variable measuring the construct of interest (BP in this instance), and a latent variable tapping the measurement bias associated with the less costly measures. As shown in **Figure 1**, the less expensive measures "load" on both of these latent variables, or factors. A key assumption of the bias control model is that the more expensive measures, by virtue of their superior construct validity, load only on the substantive (BP) factor. This latent variable, and not CBP or ABPM individually, are used to predict an outcome variable. As shown in the figure, both BP and bias factors are then allowed to be related to the substantive outcome of interest. For the present analysis, we model BP (as measured by CBP and ABPM) to predict left ventricular mass (LVM), which is an indicator of heart damage, specifically left ventricular hypertrophy.[6] We use LVM as an example, but other clinical outcomes may of course be used.

Table 1 | Breakdown of costs for clinic vs. ambulatory BP methods

	Clinic BP		Ambulatory BP	
Time for personnel (paid $25/h)	12 min	Application of cuff and three readings	5 min	Initialize monitor, enter patient data
			15 min	Validate against Korotkoff measurements
			15 min	Explanation and instructions to patient
			10 min	Download data, explain to patient
			5 min	Wash cuff, maintenance
			10 min	Other[a]
	$5 (12 min total)		$25 (60 min total)	
Equipment (amortized over 500 subjects)	$1 per participant ($500 for equipment)		$6 per participant ($3,000 per monitor)	
Subject payment	$10		$50	
Total cost	$16		$81[b]	

ABPM, ambulatory blood pressure monitoring; BP, blood pressure.
[a]Unlike clinic measurements, when using ABPM a return appointment must be scheduled, or arrangements made for the patient to return the monitor (which may cost $3,000) by mail. Some patients must be called more than once and other efforts sometimes exerted to facilitate return. We add an estimated 10 min averaged across patients. [b]We deliberately minimized the costs associated with ABPM; many researchers and clinicians would have assigned greater amounts of time and greater costs, which would only increase the rationale for using the two-method measurement design.

The present research

In this paper, we illustrate the usefulness of the TMM approach using empirical data from a study in which CBP, ABPM, and LVM measures were taken. We used the observed empirical correlations between individual measures from this study to examine hypothetical designs that are compared to a design in which CBP and ABPM measures are both taken for 300 study participants. We illustrate two outcomes of the TMM analysis, depending on the researcher's goals: (i) the benefit in statistical power, holding constant the cost of collecting CBP and ABPM data for 300 study participants; and (ii) the cost benefit due to use of the TMM design in which statistical power is held constant.

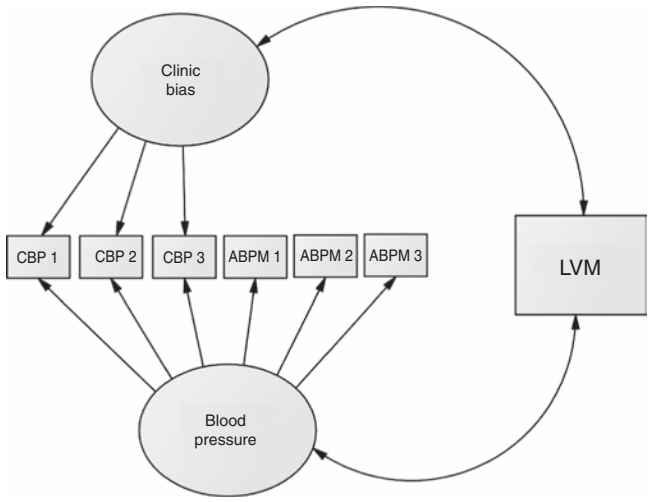

Figure 1 | Statistical model of relation of primary and secondary measures with latent (BP) construct, and prediction of LVM. ABPM, ambulatory blood pressure monitoring; BP, blood pressure; CBP, clinic blood pressure; LVM, left ventricular mass.

METHODS

We examined a data set that had been developed for other purposes, and from which other papers have been published. The methods used and rationale for the study are given elsewhere.[7] A brief synopsis of the methods is given here.

Participants. Data were available on 331 subjects (176 women, 155 men; aged 18–80; mean = 53.9), off of any hypertensive medications, with no previous cardiovascular events.

Measures, materials, and procedure. The data set contained the following variables that were germane to the present analysis:

(1) Ambulatory BP, collected using a SpaceLabs 90207 monitor (SpaceLabs, Redmond, WA), which has been shown to be valid.[8] Measurements were taken every 15 min during the day and once per hour at night for 36 h (sleep time predetermined to begin at 10:00 PM and end at 6:00 AM). These measures were repeated three times, 1 month apart. For each 36-h session, average systolic and diastolic BPs were collected for waking and sleeping hours. For the present study, we used an average of the total (waking + sleeping) systolic and diastolic BPs for each of the three 36-h periods, yielding three measures, labeled ABPM 1, ABPM 2, and ABPM 3. We also analyzed the data examining systolic and diastolic BP in separate models. We found a similar pattern of results, and thus they are not reported here.

(2) CBP was taken three times before each of the three 36-h ambulatory monitorings by a highly trained physician using a mercury column sphygmomanometer. The six BP measures (three systolic, three diastolic, one each for each 36-h ambulatory monitoring period) were averaged, yielding three measures, labeled CBP 1, CBP 2, and CBP 3. As with ambulatory BP, we found a similar pattern of results when analyzing the data using separate models for systolic and diastolic BP (not reported here).

(3) LVM was measured 10 weeks after BP data were collected via m-mode and two-dimensional echocardiograms.

Statistical procedures. The theoretical and technical aspects of the TMM design and its statistical justification as a planned missing data design have been published elsewhere.[5] The current manuscript is intended to provide an applied framework that can be used to guide the use of the method for a specific purpose, in this instance, the use of CBP and ABP. A brief overview of the TMM procedures follows.

First, we estimated the bias-correction model in the empirical data set. This model can be estimated using any structural equation modeling program with a full information maximum likelihood-based feature for handling missing data. It is useful to examine the bias of the correction model in empirical data. It is also useful to compare the key factor correlation (BP with LVM) for the bias-correction model with the same correlation when BP is defined by the CBP measures only, and when BP is defined by the ABPM measures only. Note that when the TMM design has been implemented, this first step is sufficient. However, because our purpose here is to illustrate the benefits of the TMM approach in empirical studies like this one, we took the following steps as well to explain the TMM process.

Second, the next steps require that we have a correlation matrix for the seven measures (CBP 1, CBP 2, CBP 3, ABPM 1, ABPM 2, ABPM 3, LVM). Because we must take missing data into account, we generated the correlation matrix (data not shown) using the expectation maximization algorithm.[9,10]

Third, using a non-Monte-Carlo simulation procedure (described in detail elsewhere),[5] we examined several hypothetical variations of the TMM design. For illustration purposes, we started with a "complete cases" design in which CBP and ABPM data were collected from all $N = 300$ subjects. Because the costs associated with ABPM are five times those for CBP, we can choose not to collect ABPM data for some subjects, and this allows us to collect additional data (CBP only) on many new subjects. We first explored several variations of the TMM design in which the study costs remained the same as costs for the complete cases design. Finally, we explored several variations of the TMM design for which the statistical power remained the same as the power observed for the $N = 300$ complete cases design.

RESULTS

Bias-correction model with empirical data

The bias-correction model depicted in **Figure 1** was estimated using LISREL 8.54.[11] Overall, the model fit the data well. The six BP measures had strong and significant loadings on the BP factor. The three CBP measures also had smaller, but significant loadings on the bias factor. Most important, the BP and bias factors were each significantly related to LVM, $r_{BP,LVM} = 0.43$, $r_{bias,LVM} = -0.26$. For comparison, we also tested a model in which the BP factor was defined by CBP measures only, and one in which the BP factor was defined by ABPM measures only. For the CBP-only model, $r_{BP,LVM} = 0.29$; for the ABPM-only model, $r_{BP,LVM} = 0.43$ (the difference between $r = 0.43$ and $r = 0.29$ was statistically significant, $X^2(1) = 6.98$, $P < 0.01$).

Costs constant simulation

The complete cases design had $N = 300$ with both CBP and ABPM measures. The key results for this design are shown in the top row of **Table 2**. The second row of **Table 2** shows the results for a design in which we do not need to collect ABP data for 20 of the 300 subjects. Because of the 5 to 1 cost differential between ABPM and CBP measures, the cost savings from those 20 subjects allowed us to collect CBP data for 100 new subjects, bringing the total to $N = 400$; 280 of these have data for both CBP and ABPM measures; 120 have data for the CBP only. Each subsequent row of **Table 2** shows a design in which another 20 ABPM measures were dropped, and another 100 new cases were added with CBP only. Note that all of the designs shown in **Table 2** have the same cost ($41,400) as the complete cases design.

Also shown in Table 2 is the standard error for the unstandardized factor covariance for the BP factor association with LVM. The standard error for the complete cases model (top row) was 0.055. The standard error for the design shown in the second row was actually smaller: s.e. = 0.049. In fact, as the complete cases sample size decreased, and the sample size using CBP only increased, the standard error became smaller and smaller. The standard error was smallest for the design in the eighth row of Table 2, the design with N = 170 complete cases, and N = 780 with CBP only. The statistical power associated with this design was equivalent to a complete cases design with N = 575. Thus, we would say that this design has an "effective" sample size of 575.[5]

Statistical power constant simulation

In this simulation, we also started with the complete cases design with N = 300. We then explored alternative designs, varying the numbers of ABPM and CBP measures, but keeping constant the standard error for the covariance between the BP factor and LVM. The results related to this simulation appear in Table 3. As shown in the top row of the table, the standard error for the key unstandardized factor covariance was 0.0549. In the second row of Table 3, we show the design for which we dropped the number of complete cases arbitrarily by 50. We then asked: how many subjects with CBP only must we study using this design so that the key standard error remains at 0.0549? We found the required number was 62. That is, in addition to the 50 cases for which we collected CBP data only, we must add 12 new cases (with CBP only) to retain s.e. = 0.0549 for the key factor covariance estimate. Table 3 also shows three other design alternatives. For the design in the bottom row of the table, we had 100 complete cases. With this design, we needed 351 cases with CBP only in order to retain s.e. = 0.0549 for the key factor covariance estimate.

Note in Table 3, however, that each design shown costs less than the design above it. In fact, the design shown in the bottom row of the table cost just more than half the cost of the complete cases design (top row).

Table 2 | Costs and standard error of the unstandardized factor covariance between the BP factor and LVM for varying combinations of ABPM and clinic BP measurements

Complete case	Number with only CBP	Number of subjects	Cost	s.e.	Covariance	Effective *N*
300	0	300	$41,400	0.055	0.365	300
280	120	400	$41,400	0.049		
260	240	500	$41,400	0.045		
240	360	600	$41,400	0.043		
220	480	700	$41,400	0.041		
200	600	800	$41,400	0.040		
180	720	900	$41,400	0.040		
170	780	950	$41,400	0.0396	0.367	575
160	840	1,000	$41,400	0.040		
140	960	1,100	$41,400	0.040		
120	1,080	1,200	$41,400	0.041		
100	1,200	1,300	$41,400	0.043		
80	1,320	1,400	$41,400	0.046		
60	1,440	1,500	$41,400	0.051		
40	1,560	1,600	$41,400	0.059		

The standard error shown corresponds to the unstandardized factor covariance between the BP factor and LVM. The standardized estimate, i.e., the factor correlation, rounded to r = 0.43 for both designs shown.
ABPM, ambulatory blood pressure monitoring; BP, blood pressure; CBP, clinic blood pressure; LVM, left ventricular mass.

Table 3 | Estimated costs for varying combinations of ABP and CBP measurements and resultant standard error of the unstandardized covariance between BP and LVM

Complete case	Number with only CBP	Number of subjects	Cost	s.e.	Covariance
300	0	300	$41,400	0.0549	0.365
250	62	312	$35,926	0.0549	0.365
200	129	329	$30,567	0.0549	0.365
150	212	362	$25,576	0.0549	0.366
100	351	451	$21,873	0.0549	0.367

Assumes a cost of $23 for clinic BP and $115 for ABPM.
ABPM, ambulatory blood pressure monitoring; BP, blood pressure; CBP, clinic blood pressure; LVM, left ventricular mass.

DISCUSSION

Our results showed that the standardized correlation between the BP factor and LVM was $r = 0.43$ for the bias-correction model. This correlation compared favorably with the correlation (also $r = 0.43$) observed when BP was defined using the ABPM measures only. These correlations were clearly superior (statistically and in a practical sense) to those obtained when BP was defined using the CBP measures only ($r = 0.29$). Furthermore, the magnitude of this relationship is similar to other research demonstrating that ABPM correlates with LVM in values ranging from 0.36 to 0.45.[12–14] These findings are important for the TMM design, because with this design, which is a planned missing data design, individuals are randomly selected to receive either both CBP and ABPM or just the CBP. This means that, as we showed in this work, the missing data procedures one uses to estimate the bias-correction model will have virtually the same estimates one would have if all subjects completed all measures.

In both simulations, the standardized factor correlations between BP and LVM were virtually identical in the $N = 300$ complete cases models and in all TMM models ($r = 0.43$). More importantly, using a combination of less expensive (CBP) and more expensive (ABPM) measures led to substantial improvement in statistical power (effective $N = 575$ vs. 300), as shown in **Table 2**, or a substantial cost savings for the study, as shown in **Table 3**. The results suggest the usefulness of the TMM design for two purposes: (i) to increase statistical power, while maintaining construct validity and keeping study costs constant, or (ii) to hold statistical power constant and reduce study costs.

The "TMM design" represents an advance in methods for maximizing the efficiency of a particular measure while reducing the costs of collecting gold-standard data. The present study provides insight into a specific measurement problem, ABP vs. CBP. However, the technique is applicable to any such issues that fulfill the criteria: they are related to one another, one is more expensive than the other, and the more expensive measure is considered the gold standard compared to the less expensive measure. Thus, for example, one may have to decide between intravenous blood measures (preferred) to spot measures to detect levels or increases in various neurohormones; or a structured interview compared to a paper-and-pencil measure to assess depression; or heart rate variability measures using electrocardiogram (preferred) compared to a watch that is programmed to detect heart rate. The TMM method is a promising technique that in some studies may be able to decrease costs and/or increase one's power to detect effects.

Limitations and future directions

For the assumptions of the method to hold, missing data must be planned (i.e., random assignment to ABPM+CBP or CBP-only conditions). That is, all individuals who failed to provide usable ABPM data, for example, cannot simply be grouped into a CBP-only condition, but rather the decision as to who will only have CBP data must be determined before data collection. Put another way, the missingness must be completely at random.[9] However, designs are also feasible in which the probability of receiving the more expensive measure (ABPM) is dependent on the subject's score on the less expensive measure (CBP).[5] In addition, although it is conceivable that Bayesian procedures might be developed to allow estimation of the bias control model with usual multiple regression techniques, at present, the models described here must be estimated using structural equation modeling techniques. Furthermore, the TMM approach performs best when a gold-standard measure exists. Although the TMM approach will often work well when one can simply act as though the expensive measure were a gold standard, one key element is that the scores on the expensive measure should not be under the subject's control. Also, the benefits of the TMM design are related to the cost differential between the expensive and cheap measures. We have shown here that a 5:1 cost ratio can work very well. Finally, although it may be possible in the future to adapt latent class methods to allow judgments about individual subjects, at present the TMM approach described here is best suited for research purposes and not clinical settings where judgments about individuals is paramount.

A second potential limitation concerns the cost difference to conduct the entire study, with more subjects, compared to the cost, and cost savings, for the BP measure alone. If the cost of running additional subjects is minimal, then the TMM approach is useful and recommended. If, however, the fixed costs per subject are sufficiently high (e.g., costs for assays or subject payment), they complicate assessment of the savings due to the reduced number of ambulatory monitorings that are required. Cost savings as we have described them here may be reduced in these "broad-focus" studies, and it may no longer be useful to employ a TMM design.[15] In other words, the researcher must balance the reduced cost associated with fewer ABPMs with the increased costs of running more participants. With that said, there are also hidden, or indirect benefits of the TMM approach, even in broad-focus studies, that must be taken into account when evaluating the overall benefits of the TMM approach. For example, because more participants will have completed the study, there will be greater power to detect effects for secondary analyses, such as examining how a personality characteristic such as hostility might moderate the effect of BP on LVM.

Finally, there can also be other costs associated with implementing one particular methodology over another. For example, while ABPM is considered the gold standard for BP measurement, ABPM also has lower acceptance rates by patients.[16] As with any study, researchers must weigh the benefits of collecting data using a particular methodology with the costs concerning compliance rates and data quality.

Disclosure: The authors declared no conflict of interest.

1. Verdecchia P, Schillaci G, Reboldi G, Franklin SS, Porcellati C. Different prognostic impact of 24-hour mean blood pressure and pulse pressure on stroke and coronary artery disease in essential hypertension. *Circulation* 2001; 103: 2579–2584.
2. Viera AJ, Hinderliter AL, Kshirsagar AV, Fine J, Dominik R. Reproducibility of masked hypertension in adults with untreated borderline office blood pressure: comparison of ambulatory and home monitoring. *Am J Hypertens* 2010; 23: 1190–1197.

3. Ohkubo T, Kikuya M, Metoki H, Asayama K, Obara T, Hashimoto J, Totsune K, Hoshi H, Satoh H, Imai, Y. Prognosis of "masked" hypertension and "white-coat" hypertension detected by 24-h ambulatory blood pressure monitoring. *J Am Coll Cardiol* 2005; 46:508–515.
4. Verdecchia P, Schillaci G, Borgioni C, Ciucci A, Porcellati C. White-coat hypertension: not guilty when correctly defined. *Blood Press Monit* 1998; 3:147–152.
5. Graham JW, Taylor BJ, Olchowski AE, Cumsille PE. Planned missing data designs in psychological research. *Psychol Methods* 2006; 11:323–343.
6. Verdecchia P, Schillaci G, Guerrieri M, Gatteschi C, Benemio G, Boldrini F, Porcellati C. Circadian blood pressure changes and left ventricular hypertrophy in essential hypertension. *Circulation* 1990; 81:528–536.
7. Gerin W, Ogedegbe G, Schwartz JE, Chaplin WF, Goyal T, Clemow L, Davidson KW, Burg M, Lipsky S, Kentor R, Jhalani J, Shimbo D, Pickering TG. Assessment of the white-coat effect. *J Hypertens* 2006; 24:67–74.
8. O'Brien E, Mee F, Atkins N, O'Malley K. Accuracy of the SpaceLabs 90207 determined by the British Hypertension Society protocol. *J Hypertens* 1991; 9:573–574.
9. Little RJ, Rubin DB. *Statistical Analysis with Missing Data*, 2nd edn. Wiley: New York, 2002.
10. Schafer JL. *Analysis of Incomplete Multivariate Data*. Chapman and Hall: London, 1997.
11. Jöreskog KG, Sörbom D. *LISREL 8.5 for Windows [Computer Software]*. Scientific Software International, Inc.: Lincolnwood, IL, 2001.
12. Fagard RH, Staessen JA, Thijs L. Relationships between changes in left ventricular mass and in clinic and ambulatory blood pressure in response to antihypertensive therapy. *J Hypertens* 1997; 15:1493–1502.
13. Khattar RS, Acharya DU, Kinsey C, Senior R, Lahiri A. Longitudinal association of ambulatory pulse pressure with left ventricular mass and vascular hypertrophy in essential hypertension. *J Hypertens* 1997; 15:737–743.
14. Sorof JM, Cardwell G, Franco K, Portman RJ. Ambulatory blood pressure and left ventricular mass index in hypertensive children. *Hypertension* 2002; 39: 903–908.
15. Graham JW, Shevock AE. Planned missing data design 2: two method measurement. In Graham JW (ed), *Missing Data: Analysis and Design*. Springer: New York, in press.
16. Little P, Barnett J, Barnsley L, Marjoram J, Fitzgerald-Barron A, Mant D. Comparison of agreement between different measures of blood pressure in primary care and daytime ambulatory blood pressure. *BMJ* 2002; 325:254.