Evaluation of Blood Pressure Measurement and Agreement in an Academic Health Sciences Center

Deborah S. Minor, PharmD;[1] Kenneth R. Butler, Jr, PhD;[1] Katherine L. Artman, PharmD;[2] Cathy Adair, RN, CCRC;[1] Wanmei Wang, MS;[3] Valerie McNair, LPN II;[1] Marion R. Wofford, MD, MPH;[1] Michael Griswold, PhD[3]

From the Department of Medicine, University of Mississippi Medical Center, Jackson, MS;[1] the Department of Pharmacy, University of Mississippi Medical Center, Jackson, MS;[2] and the Center of Biostatistics and Bioinformatics, University of Mississippi Medical Center, Jackson, MS[3]

The authors assessed the process of blood pressure (BP) measurement and level of adherence to recommended procedures at representative sites throughout a large academic health sciences center. A casual observer assessed the setting and observed the process, noting the equipment, technique, and BP recorded by site personnel. A trained observer then repeated the patient's BP measurement following American Heart Association recommendations. Significant biases were observed between measurements by site personnel and the trained observer. Site personnel reported on average an increased systolic BP (SBP) of 5.66 mm Hg (95% confidence interval [CI], 3.09–8.23; $P<.001$) and a decreased diastolic BP (DBP) of −2.96 mm Hg (95% CI, −5.05 to −0.87; $P=.005$). Overall, 41% of patients had a $\geq$10-mm Hg difference in SBP between measurements. Similarly, 54% had differences of $\geq$5 mm Hg in DBP between measurements. Inaccurate BP measurement and poor technique may lead to misclassification, misdiagnosis, and inappropriate medical decisions. Concordance of measured SBP between our site personnel and trained observer was less than optimal. Several areas for improvement were identified. Routine calibration and use of system-wide standardized equipment, establishment of BP measurement protocols, and periodic technique and equipment recertification can be addressed in future quality initiatives. *J Clin Hypertens (Greenwich).* 2012;14:222–227. ©2012 Wiley Periodicals, Inc.

Blood pressure (BP) measurement is perhaps the most commonly performed procedure in the clinical encounter and one of the most important measurements in clinical medicine.[1] Despite clear guidelines on appropriate techniques for BP measurement,[2,3] these recommendations rarely are followed by health care providers or personnel.[1] There are numerous factors that influence accuracy of BP readings including those related to the patient, observer, instruments, and technique.[2–4] Adequate rest time, diurnal variation, clinic atmosphere, pain, anxiety, smoking, and conversation all can have a significant impact on BP readings.[4] Factors directly related to the observer include training, end-digit preference, and impaired hearing.[2,3,5] Instrument accuracy, background noise, clothing interference, inappropriate cuff size and placement, posture, and inflation-deflation rate can influence BP measurements. Lack of repeated measurements further compounds the obtainment of an accurate reading.[3] These multiple sources of potential error encountered in daily clinical practice emphasize the possibility for inaccurate results that can influence patient management.[1]

Accurate measurement of BP is essential in staging hypertension, ascertaining BP-related risk, and guiding management. Health care providers and personnel should be keenly aware of the need to carefully follow standardized procedures in order to achieve accurate and reproducible BPs. Despite education, clinic personnel who are aware of guidelines often do not follow them to the degree necessary to produce repeatable measures, and it has been questioned whether physicians should even measure BP because they rarely follow published guidelines.[2,4,6,7] Although following guideline recommendations results in more accurate readings, health care personnel continue to use more casual methods to measure BP.[7] Using casual methods can result in several potential errors depending on the observer's knowledge or training, the equipment, and the effect of being in an ambulatory or inpatient setting.[8] Proper training of persons measuring BP and attention to such simple measures as patient positioning and appropriate selection of cuff size can increase accuracy.[2,3] The potential consequences of over- or under-treatment associated with the use of inaccurate BP measures warrant consistent practice implementation and application of the American Heart Association (AHA) guidelines.

While the gold standard for BP measurement includes the use of a mercury sphygmomanometer and identification of Korotkoff sounds, this method is no longer used in most health care settings. Accurate BP measurement can be obtained by using other methods with attention to a suitable setting, appropriately calibrated equipment, and proper technique. The level of accuracy of BP measurement throughout large academic health science centers is largely unknown. The purpose of this study was to assess the accuracy

of BP measurement and the level of adherence to recommended techniques at representative sites throughout a large academic health science center. Based on previous reports in other settings, we hypothesized that there was inconsistent use of equipment and improper BP measurement technique, thus leading to inaccurate BP measurements in our clinical and inpatient settings.

MATERIALS AND METHODS

This study was submitted and approved by our institutional review board prior to initiation. All patients seen within our health center during a 3-month period were eligible for inclusion. Patients with an arm circumference >40.9 cm were excluded. Sixteen ambulatory and inpatient sites were assessed, with the supervisor of the selected site being notified of the visit on the day of the visit. A casual observer trained in recommended BP measurement technique assessed the setting and observed the process of BP measurement by site personnel. The casual observer noted the equipment and technique used and recorded the BP obtained for each patient by the site personnel. Immediately following this, the trained observer, a certified hypertension clinical trials coordinator, repeated the patient's BP measurement according to AHA recommendations in an adjacent private area.[2,4] The trained observer used a device with a known calibration history similar to the one used by the site personnel. The BP measurements and method of assessment for each patient were recorded and provided to the health care provider.

Patient-, setting-, and technique-specific data were collected for analysis. Patient data included age, weight, arm circumference, gender, ethnicity, history of hypertension diagnosis or medications, BP, and pulse. Site and setting data included type of BP measurement device, date of last equipment calibration, type of personnel measuring the BP, and environmental factors (room temperature and set-up). Technique data included length of patient rest time, activity during BP measurement, use of bare vs covered arm, arm positioning, and patient positioning. No identifiable health information was collected for analysis or recorded.

Statistical Analysis

Means and standard deviations for continuous variables along with frequencies and percentages for categorical variables were calculated and reported. Linear mixed-effects models were used to estimate bias while adjusting for potential confounders such as patient age, weight, sex, and ethnicity and accounting for repeated measurements across patients (1 per site personnel and 2 per trained observer). A random intercept variance structure was utilized. Overall reliability was reported using intraclass correlation coefficients from the estimated mixed model covariance structure. Concordance classifications were

reported at clinically meaningful thresholds. Effect modifications in the observed bias related to deviations in specific protocol factors were examined by incorporating interaction terms in the models. Sensitivity analyses were conducted on primary model assumptions in the mean and covariance structures of the mixed models and did not alter conclusions (results available by request). Bland-Altman plots with overlaid linear regression lines were constructed to visualize potential changes in the bias and reliability metrics across level of BP.[9]

RESULTS

A total of 119 patients were observed at 16 different ambulatory or inpatient sites (57% women, 70% African American, 52% self-reported hypertension). Of the sites observed, four (25%) of the 16 were inpatient sites, representing 42 (35%) of the total (119) observations. A variety of site personnel were observed measuring BP: licensed practical nurses (LPNs, 51%), certified nursing assistants (CNAs, 33%), patient care technicians (13%), and medical doctors (MDs, 3%). Other factors assessed and observed are identified in Table I. There were 5 different BP measurement

TABLE I. Descriptive Statistics for Measurement Device, Setting, Patient Factor Observations, and Techniques Influencing Measurement Accuracy (N=119)

	No. (%)
Measurement device	
Dinamap	84 (71)
Mercury	12 (10)
Phillips	9 (7.5)
Welch-Allyn	9 (7.5)
Dinamap-antiquated	5 (4)
Measurement setting	
Central	45 (38)
Private	68 (57)
Semi-private	6 (5)
Room temperature	
Cool	4 (3)
Ambient	104 (88)
Warm	11 (9)
Observations	
Measurement of arm circumference	0 (0)
Clothing, sleeve, or garment interference	30 (25)
Out of position/bad posture	18 (15)
Improper device use	7 (6)
Stethoscope bell use (where appropriate, *n=12)	9[a] (8)
Conversation (yes)	48 (40)
Bare arm (no)	49 (41)
Arm/cuff at heart level (no)	89 (75)
Feet not flat (no)	77 (65)
Crossed legs (yes)	24 (20)
Second blood pressure measurement (no)	114 (95)
Clinic rest time (no)	117 (98)

TABLE II. Comparison of Mean±SD (mm Hg) BP Measures of Site Personnel and Trained Observer and Concordance

	Site Personnel	Trained Observer	Bias (Average Difference)		Concordance/Reliability						
					Absolute Differences (mm Hg)						
					≥ 5	≥ 10	≥ 15	≥ 20	≥ 25	≥ 30	ICC
SBP	129 (125–133)	125 (122–128)	Unadjusted	3.86 (1.67–6.05) *P*=.001	71%	41%	24%	16%	5%	1%	0.87
			Adjusted[a]	5.66 (3.09–8.23) *P*<.001							(0.82–0.91)
DBP	74 (72–76)	76 (75–78)	Unadjusted	−2.27 (−3.73 to −0.81) *P*=.002	54%	21%	8%	4%	0%	0%	0.78
			Adjusted[a]	−2.96 (−5.05 to −0.87) *P*=.005							(0.70–0.84)

Abbreviations: DBP, diastolic blood pressure; ICC, intraclass correlation coefficient; SBP, systolic blood pressure. [a]Adjusted for patient age, weight, sex, and ethnicity.

devices used at the sites, with the Dinamap (GE Healthcare, Waukesha, WI) being the most frequent (71%). A calibration history was available at only the site using mercury sphygmomanometers. Most observed BPs were measured in a private setting (57%). There was no rest period between arrival and BP measurement for 98% of patients; measurement of arm circumference prior to cuff application was not observed in any patient. A second BP was not obtained by site personnel in 95% of observations. The stethoscope bell was utilized in 9 of 12 measurements requiring this technique. The casual observer noted the following technical parameters: conversation during measurement (40%), measurement taken on a clothed arm (41%), cuff and arm not positioned at heart level (75%), feet not flat on the floor (65%), and legs crossed (20%).

Site personnel and trained observer BP measurements were compared using unadjusted and adjusted average differences for SBP and DBP as well as concordance and reliability measures (Table II). The adjusted average difference for SBP was 5.66 mm Hg (95% confidence interval [CI], 3.09–8.23; *P*<.001), indicating that site personnel tended to report higher SBPs. For DBP, the adjusted difference was −2.96 mm Hg (95% CI, −5.05 to −0.87; *P*=.005), indicating that site personnel tended to report lower DBPs on average. An SBP difference ≥10 mm Hg was observed in 41% of patients, while a DBP difference ≥5 mm Hg was observed in 54% of patients (Table II). Although intraclass correlation coefficients (ICCs) indicated that overall associations between measurers were high (SBP: *r*=0.87; 95% CI, 0.82–0.91 [*P*<.001]; DBP: *r*=0.78; 95% CI, 0.70–0.84 [*P*<.001]) by this method of analysis, the more relevant the poor absolute agreement as indicated by the Bland-Altman method (Figure 1). Using this method, SBP readings were ±10 mm Hg in 59% of observations, while DBP readings were within ±5 mm Hg in 46% of observations. Compared with these findings, the high ICC scores for both SBP and DBP indicating high levels of agreement are misleading. Contrary to the indication of the ICC, when examining the incremental observed

differences using absolute agreement and the Bland-Altman method, there is much variability and poor overall agreement between site personnel and trained observers for both SBP and DBP.

Cumulative effect and interference factors were grouped in an ordinal fashion by adding the number of observations noted based on type of setting, room temperature, patient factors, or technique, which could incrementally interfere with the accuracy of the reading and better account for differences in BP measurements between site personnel and the trained observer. An overwhelming majority of participants had ≥1 factors noted that would influence the accuracy of the BP measurement obtained by site personnel (98%).

Some environmental, patient, and technique factors potentially contributed to the magnitude of the site personnel vs trained observer differences (Table III). Site personnel reported SBPs only 1.68 mm Hg higher if the patient had a bare arm, but 6.56 mm Hg higher when the patient did not have a bare arm (unadjusted effect modification, 4.88 mm Hg; 95% CI, 0.53–9.23 [*P*=.028]). Thus, if all site personnel had followed the bare-arm protocol, the difference in SBP measures may have been less substantial. We note that the bare-arm effect modification was diminished but similar in the adjusted model. Other effect modifiers analyzed including feet not flat, conversation, cuff position (arm not at heart level), and crossed legs contributed to mild to modest variability in readings (Figure 2); however, none were statistically significant in either adjusted or unadjusted models. Differences for DBP were even smaller, and none were statistically significant in either model. We also conducted a sensitivity analysis comparing ambulatory vs inpatient site types and estimates did not change (Table SI). Several important factors were not included in the effect modification analyses because of low numbers of observations, making it impossible to determine differences with and without the effect: inadequate rest time (all but 2 patients had <1 minute rest time), arm circumference measurement (0 for site personnel), and nonambient temperatures (11).
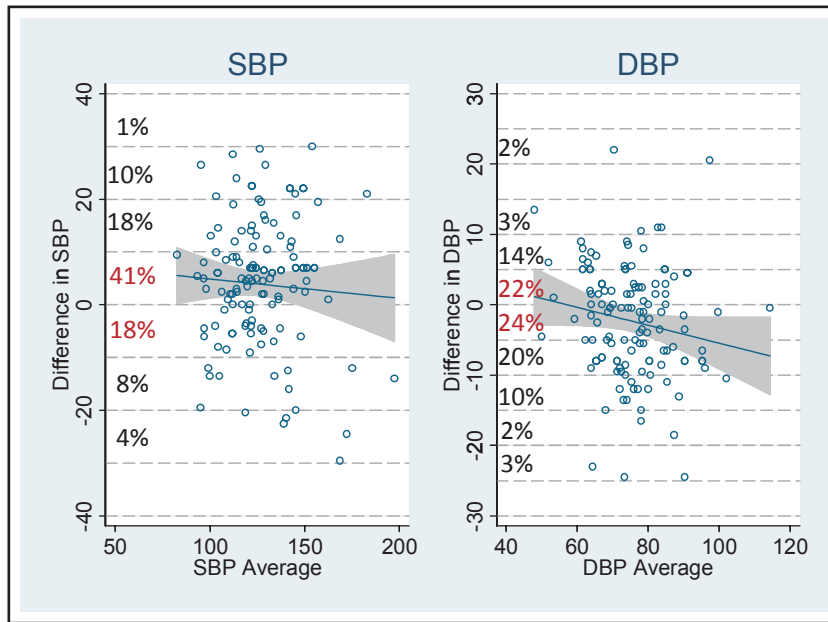
FIGURE 1. Graphical representation of observed blood pressure (BP) differences using Bland-Altman method plotting differences in BP of site personnel and trained observer against the average BP. For systolic BP (SBP), there was agreement within ±10 mm Hg in 70 observations (59%) between site personnel and the trained observer. For diastolic BP (DBP), there was agreement in 72 (46%) of observations that fell within ±5 mm Hg.

TABLE III. Effect With and Without Modifiers on (A) SBP and (B) DBP Readings by Site Personnel and Trained Observer With Bias

Effect Modification	Difference Between Site Personnel and Trained Observer		Effect Modification (Bias)	
Effect modifier	With effect (not per-protocol)	Without effect (per protocol)	Unadjusted	Adjusted
(A)				
Non–bare arm (sleeve)	6.56 (3.21–9.91)	1.68 (−1.10 to 4.45)	4.88 (0.53–9.23) P=.028	4.25 (−0.85 to 9.35) P=.103
Feet not flat	3.42 (0.13–6.70)	6.95 (3.44–10.46)	−3.54 (−8.34 to 1.27) P=.149	−2.82 (−8.13 to 2.50) P=.299
Conversation	5.99 (2.62–9.36)	2.22 (−0.80 to 5.24)	3.77 (−0.75 to 8.29) P=.102	3.31 (−1.94 to 8.57) P=.216
Arm/cuff not at heart level	6.21 (1.52–10.90)	3.58 (1.05–6.12)	2.62 (−2.71 to 7.96) P=.335	1.32 (−5.83 to 8.47) P=.717
Legs crossed	3.39 (−1.86 to 8.65)	4.10 (1.63–6.57)	−0.71 (−6.51, 5.10) P=.811	−0.47 (−7.74 to 6.80) P=.899
(B)				
Non–bare arm (sleeve)	−2.25 (−4.57 to 0.07)	−2.25 (−4.17 to −0.33)	−0.004 (−3.02 to 3.01) P=.998	1.41 (−2.73 to 5.55) P=.504
Feet not flat	−1.63 (−4.04 to 0.79)	−3.67 (−6.24 to −1.09)	2.04 (−1.49, 5.57) P=.261	2.56 (−1.66 to 6.79) P=.231
Conversation	−1.86 (−4.21 to 0.48)	−2.31 (−4.41 to −0.21)	0.45 (−2.71 to 3.60) P=.781	−0.12 (−4.30 to 4.05) P=.954
Arm/cuff not at heart level	−2.01 (−5.18 to 1.17)	−2.25 (−3.96 to −0.54)	0.24 (−3.37 to 3.84) P=.897	−1.87 (−7.50 to3.76) P=.526
Legs crossed	−0.14 (−3.65 to 3.37)	−2.83 (−4.47 to −1.18)	2.69 (−1.19 to 6.56) P=.174	2.77 (−2.97 to 8.51) P=.345

Abbreviations: DBP, diastolic blood pressure; SBP, systolic blood pressure. Factors that were not included in the effect modification analyses: inadequate rest time (all but 2 patients had <1 minute rest time), arm circumference measurement (0 for site personnel), and nonambient temperatures (11).

DISCUSSION

To our knowledge, this is the first study to examine both the process of BP measurement and the level of agreement between repeated measures in a large academic health sciences center. This study supports that AHA guidelines are not generally followed in most of the sites we observed, with significant variability and poor overall agreement demonstrated between site per-

sonnel and trained observer measures for both SBP and DBP. Numerous factors influence the level of agreement and reproducibility of BP measurements, including technique, accuracy of devices, environmental setting, and patient factors.[2,10–13] All are critical and ultimately influence accuracy and repeatable results. Pickering and colleagues[2] demonstrated that most devices reviewed have inherent rates of
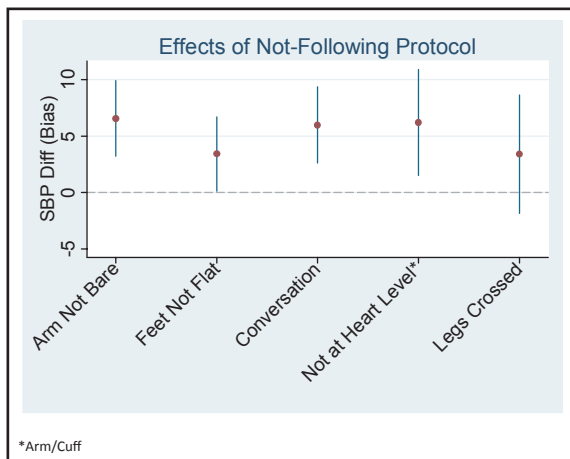
FIGURE 2. Diagram highlighting conclusions of bias for individual protocol violations. Arm not being bare, conversation, and cuff not at heart level demonstrated the greatest mean differences of >5 mm Hg. Feet not flat and legs crossed did not appear to influence systolic blood pressure (SBP) differences (diff) as much as other variables.

inaccuracy. Device accuracy is difficult to determine without documentation of calibration history. Inaccurate BP measurements can have significant connotations particularly in patients with BPs at or near diagnostic or treatment thresholds. These patients could easily be misclassified and inappropriately managed as a result of inaccurate assessment. In a recent study, Powers and colleagues[14] suggested that within-patient variability could be greatly reduced by averaging 5 or 6 measurements with additional measures for decision-making in patients who are closer to treatment thresholds. We found that although patient and environmental factors did cause some variability in BP readings, they accounted little for statistically different variations between site personnel and trained observer readings, suggesting that most of the variation may lie in the devices used or differences in technique.

Perhaps the most interesting findings in this study was the high correlation (as indicated by the ICC) of both SBP and DBP measures compared with the poor agreement when comparing thresholds at 5 mm Hg and 10 mm Hg using the Bland-Altman method. The ICC identified that correlations between trained observer and site personnel readings were similar, particularly in the case of high BP readings. Although this test revealed good agreement between the trained observer and site personnel BP readings, this single measure of correlation is misleading. The analysis was therefore extended using the Bland-Altman method with a review of differences in readings of 10 mm Hg and 5 mm Hg for SBP and DBP, respectively. These observed differences more completely describe the level of agreement and are more clinically useful.

Most sites have transitioned away from standard mercury sphygmomanometers and now use automated devices. The lack of health system standardization of equipment and routine calibration of these devices is problematic. Moving to a single type of equipment and measurement across the health system has the potential to improve measurement variability. Site personnel could receive consistent training on device use with the establishment of core competencies and periodic re-certification. However, standardization of equipment and redesigning clinical environments would be a substantial economic investment for any health care system. Quality initiatives regarding recommended guidelines for BP measurement and simple modifications of less-than-optimal clinic or site conditions are practical and more easily accomplished. Academic training programs for all health care providers should also include training in accurate BP measurement.

STRENGTHS AND LIMITATIONS
Strengths of this study include the use of statistical methods to demonstrate differences in measurements rather than general summary measures and standard correlations. The use of experienced and certified hypertension clinical trial coordinators gives us confidence in the repeated BP measures. We were able to identify several areas for quality improvement to increase the accuracy of BP readings across our institution. Limitations to this study include the use of aggregate data from a convenience sample of both observations and sites. This could have resulted in selection bias, although we attempted to include a sampling from as many clinics and inpatient settings as possible. Several important factors including inadequate rest time, arm circumference measurement, and nonambient room temperatures were not included in the effect modification analyses because of a low number of observations, making it impossible to determine differences with and without the effect. In addition, BP is inherently variable; therefore, some degree of variability is always expected between readings on the same patient.

CONCLUSIONS
Inaccurate BP measurement and poor technique can lead to misclassification, misdiagnosis, and inappropriate medical decisions. From a clinical perspective, agreement of BP measures between site personnel and the trained observer in this study was poor. This study revealed several areas for improvement in BP measurement in our academic health sciences center and the need for future quality initiatives. The analytic approach we employed demonstrated that reporting of summary statistics alone can be misleading in regard to the comparison of BP measurement. Our findings are likely not unique and probably reflect common and universal problems throughout health care environments. These findings simply highlight the increased vigilance needed to ensure the accurate assessment of BP and the role academic health centers should have in addressing this issue and the training of others.

References

1. Pater C. Beyond the evidence of the new hypertension guidelines. Blood pressure measurement – is it good enough for accurate diagnosis of hypertension? Time might be in, for a paradigm shift (I). *Curr Control Trials Cardiovasc Med*. 2005;6:6.
2. Pickering TG, Hall JE, Appel LJ, et al. Recommendations for blood pressure measurement in humans and experimental animals, part 1: blood pressure measurement in humans: a statement for professionals from the Subcommittee of Professional and Public Education of the American Heart Association Council on High Blood Pressure Research. *Hypertension*. 2005;45:142–161.
3. Pickering TG, Hall JE, Appel LJ, et al. Recommendations for blood pressure measurement in humans: an AHA scientific statement from the Council on High Blood Pressure Research Professional and Public Education Subcommittee. *J Clin Hypertens*. 2005;7:102–109.
4. Chobanian AV, Bakris GL, Black HR, et al. Seventh report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. *Hypertension*. 2003;42:1206–1252.
5. Pickering TG. Principles and techniques of blood pressure measurement. *Cardiol Clin*. 2002;20:207–223.
6. Graves JW, Sheps SG. Does evidence-based medicine suggest that physicians should not be measuring blood pressure in the hypertensive patient? *Am J Hypertens*. 2004;17:354–360.
7. Armstrong RS. Nurses' knowledge of error in blood pressure measurement technique. *Int J Nurs Pract*. 2002;8:118–126.
8. Dickson BK, Hajjar I. Blood pressure measurement education and evaluation program improves measurement accuracy in community-based nurses: a pilot study. *J Am Acad Nurse Pract*. 2007;19:93–102.
9. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1:307–310.
10. Intrumentation AftAoM. *American National Standard for Electronic and Automated Devices* 1993. ANSI/AAMI SP10-1992.
11. White WB, Berson AS, Robbins C, et al. National standard for measurement of resting and ambulatory blood pressures with automated sphygmomanometers. *Hypertension*. 1993;21:504–509.
12. Parati G, Stergiou G. Self blood pressure measurement at home: how many times? *J Hypertens*. 2004;22:1075–1079.
13. Vinyoles E, Blancafort X, Lopez-Quinones C, et al. Blood pressure measurement in an ambulatory setting: concordance between physician and patient self-measurement. *J Hum Hypertens*. 2003;17:45–50.
14. Powers BJ, Olsen MK, Smith VA, et al. Measuring blood pressure for decision making and quality reporting: where and how many measures? *Ann Inter Med*. 2011;154:781–88.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table SI. Protocol violation cross-classification frequencies and percentages.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.