National Health Statistics Reports

Number 59 ■ October 5, 2012

Blood Pressure Randomized Methodology Study Comparing Automatic Oscillometric and Mercury Sphygmomanometer Devices: National Health and Nutrition Examination Survey, 2009–2010

by Yechiam Ostchega, Ph.D., R.N., Division of Health and Nutrition Examination Surveys; Guangyu Zhang, Ph.D., Office of Research and Methodology; Paul Sorlie, Ph.D., National Institutes of Health, National Heart, Lung, and Blood Institute; and Jeffery P. Hughes, M.P.H.; Debra S. Reed-Gillette, M.S.; Tatiana Nwankwo, M.S.; and Sarah Yoon, Ph.D., R.N., Division of Health and Nutrition Examination Surveys

Abstract

Objectives—The mercury sphygmomanometer has been the gold standard used for obtaining blood pressure (BP) for the National Health and Nutrition Examination Survey (NHANES) from 1960 to the present. However, due to environmental concerns and an increased use of automated oscillometric BP devices, NHANES has been exploring an alternative to using the standard mercury sphygmomanometer (mercury) to measure BP.

Methods—The accuracy of Omron HEM–907XL BP readings was compared with that of mercury BP device readings for gender, age group, race and ethnicity, and body mass index categories and cuff-size subgroups. Each person had three BP measurements per device recorded sequentially. The order of the devices and readers were randomly assigned. A total of 6,460 participants had three valid systolic readings, and 6,338 had three valid diastolic readings.

Results—Omron and mercury measurements were correlated ($r = 0.92$, systolic BP; $r = 0.79$, diastolic BP). Overall, the mean between-device differences (Omron and mercury) were –1.6 mm Hg for systolic and –0.6 mm Hg for diastolic ($p < 0.05$ for both). The mean between-device differences were less than or about 2 mm Hg for each subgroup: gender, age group, race and ethnicity, and body mass index categories, and cuff-size subgroups. The exceptions were mean systolic between-device differences for those using the extra-large BP cuff (–3.1 mm Hg) and obese individuals (–2.6 mm Hg), and the mean diastolic between-device differences for the underweight group (–3.5 mm Hg). Assuming mercury to be the gold standard, between-device agreements for the frequency of high BP (140/90 mm Hg or more) and stage II high BP (160/100 mm Hg or more) were above chance ($\kappa = 0.72$ for both). Omron underestimated the high BP frequency by 2.28% and stage II high BP frequency by 0.77%.

Conclusions—Lower estimates of high BP by the Omron device may require adjusting future national prevalence estimates accordingly to account for between-device differences.

Keywords: National Health and Nutrition Examination Survey • mercury oscillometric devices

Introduction

For many years, the standard instrument for measuring blood pressure (BP) in the National Health and Nutrition Examination Survey (NHANES) has been the mercury sphygmomanometer. In recent years, because of increased environmental concerns about the disposal of mercury-contaminated medical waste and the risk of spills from mercury sphygmomanometers, clinical settings (e.g., doctor's offices, clinics, and hospitals) have begun phasing out the mercury devices. In addition, in a number of cross-sectional state surveys (1–4), an automatic oscillometric device has replaced the mercury device in assessing BP. In 1999, the Environmental Protection Agency and the American Hospital Association published a report in which both agencies agreed to eliminate mercury-containing waste from the health care industry by 2005 (5). In that report, both agencies' recommended best practices were discussed and, based on those recommendations, federal hospitals eliminated mercury-containing instruments, including

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Centers for Disease Control and Prevention
National Center for Health Statistics

sphygmomanometers. Similarly, the National Institutes of Health's ''Mad as a Hatter'' Mercury-Free Campaign goal is to eliminate all unnecessary uses of mercury in NIH facilities (6). Along these lines, NHANES has been considering alternatives to the mercury sphygmomanometer for future survey cycles.

Among the automatic oscillometric devices, the Omron BP monitor is designed and utilized as a clinical BP monitor and was used by the National Heart, Lung, and Blood Institute study known as Coronary Artery Risk Development in Young Adults or CARDIA (7). Omron was previously validated in multiple studies (8,9); all past validation studies were performed using small sample sizes and in a study setting. Therefore, in 2006–2007, NHANES conducted a study using a convenience sample of 509 persons in the mobile examination center (MEC) that compared the mercury sphygmomanometer and the Omron HEM–907XL BP monitor, following the Association for the Advancement of Medical Instrumentation's (AAMI) 2002 criteria as a guideline (10,11). With the exception of diastolic BP in youths aged 13–19 years [mean difference –1.77 mm Hg; standard deviation (SD), 8.65], the Omron device met the AAMI validation criteria (11).

In contrast with the first study, which used a convenience sample, the motivation for the second study was to determine whether the results of the first study could be replicated in the usual NHANES MEC standardized data collection environment with regular survey participants. Therefore, in early 2009, a BP randomized methodology study was conducted during a 2-year NHANES data collection cycle (2009–2010) to compare the Omron and mercury devices with the ultimate goal of assessing whether the mercury sphygmomanometer can be replaced with the Omron device,

The overall objective of this study was to compare mercury sphygmomanometer (mercury) readings to Omron HEM–907XL (Omron) readings using a randomized study design in a standard NHANES environment. The specific objectives of the analysis were twofold. The first objective was to compare the measured values of BP obtained by Omron with those obtained by mercury overall and by gender, age, race and ethnicity, BP cuff size, body mass index (BMI), and irregular heart rate. The second objective was to compare high BP classification agreement between the devices.

Methods

Study population

The 2009–2010 BP randomization study participants were from the 2009–2010 NHANES. NHANES is a cross-sectional national health survey of the civilian, noninstitutionalized U.S. population, conducted by the Centers for Disease Control and Prevention's National Center for Health Statistics (NCHS). Descriptions of the sample design and data collection methods for NHANES are available on the survey website (12). Survey participants were interviewed in their homes and then examined in the NHANES MEC. The BP randomization study was part of the NHANES MEC data collection cycle for 2009–2010. The NCHS Ethics Review Board approved the BP randomization study protocol.

Sample selection

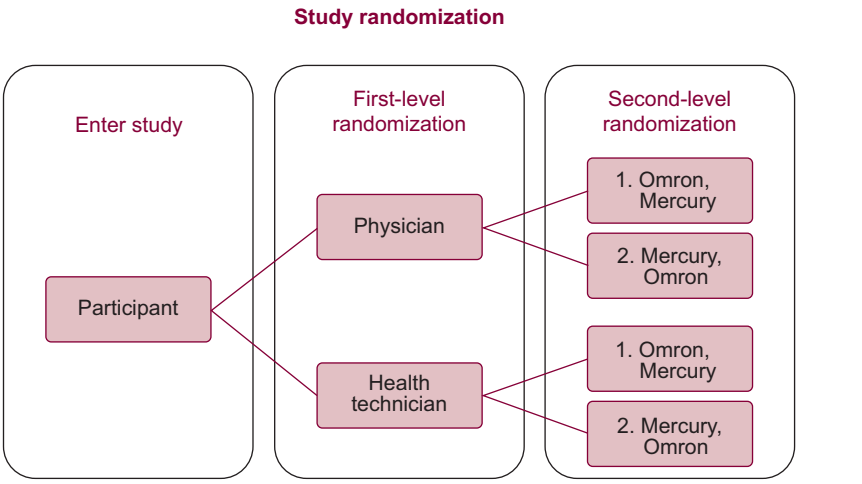
All of the participants in the BP randomization study were regular NHANES MEC participants. The BP randomization study started approximately 2½ months after the beginning of the 2009 NHANES data collection cycle, consequently missing three data collection sites. The late start date was related to logistic and training issues. A total of 7,410 participants aged 8 years and over were eligible to participate in the BP randomization study. Of these, 343 were excluded due to having all BP values missing, and 7 were excluded because an observer made too few measurements. For analysis of systolic BP, an additional 722 individuals were excluded due to having fewer than three systolic readings, leaving 3,231 males and 3,229 females. For analysis of diastolic BP, an additional 600 were excluded due to having fewer than three diastolic readings, leaving 3,166 males and 3,172 females (Table 1).

Following AAMI standards, the term ''observer'' is used throughout this report to denote a health technician or physician trained to accurately obtain BP values using the mercury device (10,11). A total of 6,460 persons aged 8 years and over were eligible to participate in the BP study. Participants had three valid mercury and Omron systolic readings (3,231 males and 3,229 females) and 6,338 persons had three valid mercury and Omron diastolic readings (3,166 males and 3,172 females). For more details about the sample selection, see Table 1.

The overall mean age of the study participants was 40 years (age range 8–80 years and over). No statistically significant difference was observed between participants and eligible nonparticipants across age (40.2 compared with 41.6 years, $p = 0.05$), BMI (27.3 compared with 27.7, $p = 0.05$), and midarm circumference (31.6 compared with 31.5 cm, $p \geq 0.05$). When compared with males, a greater percentage of females did not participate in the study (males 9.52% compared with females 11.45%, $p < 0.01$). A greater percentage of non-Hispanic black persons did not participate in the study compared with non-Hispanic white persons, Hispanic persons, and other racial and ethnic categories (non-Hispanic black, 12.87%; non-Hispanic white, 9.92%; Hispanic, 9.54%; and other, 11.52%; $p < 0.01$).

Study design

Study participants were randomly assigned to have their BP taken either by a MEC physician or by a health technician. Next, the order of which device was to be used initially was further assigned (i.e., each person randomly received an Omron reading or mercury reading first) (Figure 1). When using the Omron, the ''hide'' mode feature was utilized. Once measurements



SOURCE: CDC/NCHS, National Health and Nutrition Examination Survey, 2009–2010 blood pressure methodology study.

Figure 1. Randomized sampling methodology for blood pressure study comparing Omron and mercury devices: National Health and Nutrition Examination Survey, 2009–2010

were captured, the three readings were recorded by a staff member, who did not know the readings from the other instrument. This specific design approach was taken for three reasons: first, to minimize observer bias; second, to minimize the order effect (i.e., first reading being always higher) (13); and third, to minimize the possibility that Omron readings may affect the mercury readings. Therefore, BP values were keyed off the Omron device by persons who were effectively blind to the mercury results.

The data confirmed that the randomization schema worked well, with 49% of the readings being done by a health technician and 51% being done by a physician. Additionally, 50% of participants had their BP measured initially with the Omron followed by the mercury, and the other 50% had the mercury followed by the Omron. The slightly greater percentage of physicians than technicians taking a BP reading was related to the fact that the physicians started taking study BP measurements approximately 1 month earlier than the technicians. The physicians were trained and started the randomized study 2½ months after the start of the 2-year cycle, whereas the technicians started collecting data 1½ months later. Therefore, the first 1½ months of the study had only one randomization schema, namely device order. The random order of measurements (mercury first or Omron first) resulted in equal means of age, BMI, and arm circumference in the two groups ($p > 0.05$ for all). Additionally, the distribution of gender, BMI, race and ethnicity, and cuff-size groups also showed equal distribution between the randomized orders ($p > 0.05$ for all) (14).

Equipment

The Omron HEM–907XL is a digital upper-arm electronic blood pressure monitor that is designed to be used in clinical settings (15). According to the manufacturer (15), the HEM–907XL is an updated version of the HEM–907, specifically because it has been upgraded to allow inflation of a larger-size cuff (15). The algorithm range was expanded to accept the measurement for a larger arm circumference while still maintaining its validated margin of error (± 0.3%). The Omron HEM–907XL automatic measurements are based on smart "inflate" technology (IntelliSense), where inflation is driven by a pumping system, and deflation is driven by an electromagnetic control valve that allows rapid air release. The measurement scale for this oscillometric device ranged from zero to 280 mm Hg. A special function "hide" mode conceals blood pressure values from a person's view, which may reduce participant anxiety (15).

The medical wall-mounted mercury sphygmomanometer (Baumanometer) was used as the standard comparison device to the Omron HEM–907XL. The study employed only one brand of cuffs: the standard mercury Baumanometer cuffs [small adult (17–22 cm), adult (22–32 cm), large adult (32–42 cm), and extra-large adult (42–50 cm)] were used to perform all blood pressure measurements with both the mercury sphygmomanometer and the Omron HEM–907XL.

The calibration technique performed to check the HEM–907XL involved connecting the mercury sphygmomanometer and the Omron device via a T-tube connector and setting the Omron HEM–907XL MODE selector to "CHECK." Explicitly, the pressure values displayed on the monitor were compared with the ones on the mercury manometer at different BP points (300, 250, 200, 150, 100, 60, and 0 mm Hg) using the four Baumanometer cuff sizes wrapped around a cylinder. According to the manufacturer's guidelines, comparative readings needed to be within ± 3 mm Hg to meet the calibration procedure criteria (15).

Measurement of BP

Performing the BP readings required participants to be seated in a chair with back support, with both feet resting comfortably on the floor, and both forearms supported on a level surface at heart level. The appropriate BP cuff size (small adult, adult, large adult, extra-large adult) was selected according to the midarm circumference of the participant (16). After a 5-minute rest, the study participants had their BP measured. Three systolic and diastolic measurements were obtained by the first assigned device, and the next three systolic and diastolic measurements were obtained by the second device. For mercury measurements, a fourth reading was taken if one or two of the other

readings were not obtained. Both the mercury individual determination (systolic/diastolic) and Omron individual determination (systolic/diastolic) were repeated at 30-second intervals. The transition from one device BP determination to the other also was obtained within 30 seconds.

Fourteen observers were involved in the BP randomization study (10 health technicians and 4 physicians). Fifty-one percent of all BP readings were done by the physicians; the majority of those BP readings (90%) were done by two of the four physicians. The mean number of readings per observer was 462, median 293, minimum 35, and maximum 1,836. Three criteria were used to assess the observer effect on the BP readings of the devices: a) individual observer's mean difference of the between-device readings, b) end-digit preference, and c) BP cuff selection. All criteria assessing observer effect were based on our previous work assessing BP quality assurance in NHANES (17). The 14 BP observers' between-device readings did not substantially deviate from each other; 75% of between-observer differences were within 3 mm Hg for systolic, and 70% were within 3 mm Hg for diastolic (differences are in absolute values). The between-device values by individual observer for systolic BP reading ranged from a mean of –3.4 mm Hg to 2.7 mm Hg, and for diastolic BP reading, the values ranged from –3.6 mm Hg to 3 mm Hg.

For the four mercury systolic BP determinations, all of the end-digits were at or about 20% preferences. As for mercury diastolic BP values, the fourth determination had the largest end-digit preference for zero (27%). The fourth reading was always taken if one or two of the other readings were not obtained. Therefore, this reading was the most difficult to obtain, which could account for the greater than 20% end-digit preference for zero. After rounding the Omron readings to even end-digit values, the end-digit preferences of the three Omron systolic and diastolic readings were analyzed. All of the end-digit preferences were in the 19%–20% range.

The observers' cuff selections were compared with midarm circumference values obtained during the body measurements exam in the MEC. A total of 541 persons, or 8%, had discrepant cuff size. Discrepant cuff size is defined as an observer selecting a BP cuff that is different from a cuff size selected with the body measure component of midarm circumference values. The discrepancy resulted in more participants with a larger cuff than needed rather than a smaller cuff. Specifically, among adult cuff selections, 6.13% were overcuffed and given a large adult cuff size, while only 0.48% were undercuffed and given a small adult cuff size. Among large adult cuff selections, 2.81% were overcuffed and given an extra-large adult cuff size, and 1.81% were undercuffed and given an adult cuff size. The above notwithstanding, our findings show that 92.56% of the small adult/child BP cuffs, 93.39% of the adult BP cuffs, 95.30% of the large adult BP cuffs, and 94.53% of the extra-large BP cuffs were correctly selected. Only three individuals required an infant cuff.

Other measurements

Height in centimeters and weight in kilograms were measured in the MEC following a standard protocol. BMI was calculated as weight in kilograms over height in meters squared (kg/m^2). Midarm circumference was determined by having the participant stand erect with his or her feet together, his or her right arm flexed at a 90° angle at the elbow, and his or her palm facing up. On the right scapula, the observer located and marked, with a horizontal line, the uppermost edge of the posterior border of the acromion process. The observer held the zero end of the measuring tape at this mark and extended the tape down the posterior surface of the arm to the tip of the olecranon process. The observer made a horizontal mark at the midpoint at the posterior aspect of the arm and measured the arm circumference.

Statistical analyses

All statistical analyses were performed using the software products SAS 9.2 for Windows (Cary, N.C.) and STATA 11.1 (College Station, Texas).

Assessing device differences

The statistics used to assess device differences were based on an AAMI recommendation for noninvasive sphygmomanometers—clinical validation of automated measurement devices—and previous validation studies (11,12). For each person, the average of the three measurements was calculated from the same device, which leads to four readings per subject denoted as systolic mercury, systolic Omron, diastolic mercury, and diastolic Omron. The between-device differences for systolic BP and diastolic BP were assessed separately. All of the differences were calculated as Omron–mercury. The correlation of BP readings between the two devices was calculated using the Pearson correlation coefficient. A graphic display of the differences of between-device readings (Omron – mercury) compared with the corresponding averages [(Omron + mercury/2)], a Bland and Altman graph, was created to assess the relationship between the two devices (18). The Bland and Altman scatter-plot graphic displays were overlaid with a local regression line (LOESS) to further assess the relationship between the two devices. More specifically, LOESS was used as a smoothing algorithm to allow greater flexibility to fit a regression line in the face of possible outliers in the data (19).

In addition, absolute between-device differences were calculated and displayed graphically, following both the AAMI and the international European protocol for device comparison (11,20). For more details about differences between devices, we also compared BP values across devices by selected percentiles (1, 10, 25, 50, 75, 90, 95, and 99%).

The mean difference (Omron minus mercury) of the test device (Omron) and

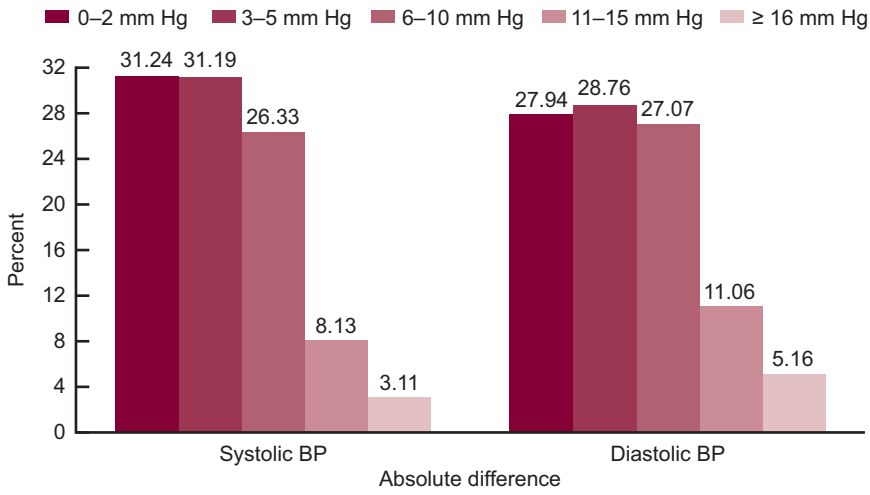
the comparison device (mercury) were calculated within strata:

1. By gender.
2. By four age groups (8–18, 19–39, 40–59, and 60 and over).
3. For three race-and-ethnicity self-reported categories (Hispanic, non-Hispanic white, and non-Hispanic black).
4. For four cuff sizes (child/small adult, adult, large adult, extra-large adult).
5. For four BMI categories in persons aged 18 and over [underweight (less than 18.5 kg/m^2), normal weight (less than 25.0 kg/m^2), overweight (25.0 to 29.9 kg/m^2), and obese (30 kg/m^2 or more)].
6. By presence of an irregular heart rate.

Paired *T* tests were used, separately, to test if the mean between-device differences were equal to zero for systolic and diastolic across the categories listed above. The α-level for a significant test was considered to be $p < 0.05$.

Comparing high BP classification agreement

As in our previous work, we used agreement statistics to assess between-device agreement for high BP classification (12). Survey participants aged 18 and over who had three valid mercury and Omron systolic readings and three valid mercury and Omron diastolic readings were included in the analysis. Classifications of high BP (i.e., BP greater than or equal to 140/90 mm Hg) and stage II high BP (i.e., BP greater than or equal to 160/100 mm Hg) were compared based on mercury and Omron readings (21). Note that high BP was defined by BP values only, regardless of whether the study participant took antihypertensive medication. The agreement parameters estimated were sensitivity, specificity, and kappa agreement. Mercury was considered the gold standard.



NOTE: BP is blood pressure in mm Hg.
SOURCE: CDC/NCHS, National Health and Nutrition Examination Survey, 2009–2010 blood pressure methodology study.

Figure 2. Frequency distribution of absolute between-device differences (Omron minus mercury) for systolic and diastolic blood pressure values

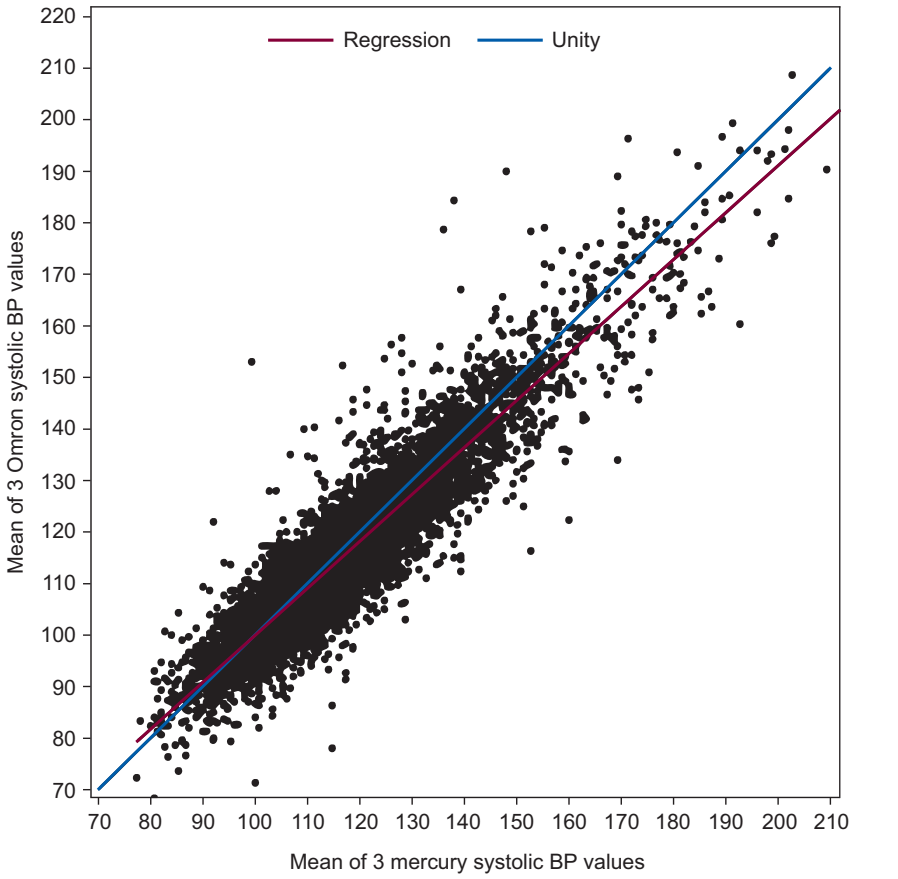
Results

Comparing measured values of BP (Omron – mercury)

Figure 2 is a bar graph that shows the percent distribution of the absolute differences between the two device measurements within 0–2, 3–5, 6–10, 11–15, and 16+ mm Hg categories. Absolute agreement within 5 mm Hg is considered the acceptable threshold for between-device agreement (11). In our study, the absolute between-device differences were within 5 mm Hg for 62.43% of the systolic BP readings and 56.71% for diastolic BP readings. The absolute agreement within 2 mm Hg was 31.24% for systolic readings and 27.72% for diastolic readings.

Figures 3 and 4 show scatter plots of the mercury and Omron BP readings, averaged over the three measurements, for systolic and diastolic separately. The respective correlations were statistically significant ($r = 0.92$ systolic, $r = 0.79$ diastolic; $p < 0.001$ for both). The scatter plots were overlaid by a unity and a regression line. In general, the regression line was below the unity line for systolic BP, suggesting that Omron underestimated the mean of systolic BP compared with mercury; for diastolic BP, the regression line was above the unity line for the lower BP values and was below the unity line for higher BP values, suggesting that Omron overestimated the mean of BP at low diastolic mercury values and underestimated the mean of BP at high mercury diastolic values. Table 2 compares selected percentile values by device. For systolic BP, mercury always read higher than Omron. In contrast, Omron read higher than mercury for lower diastolic BP (1%–10%), and read lower than mercury for higher diastolic BP (25% or more).

Figures 5 and 6 show the relationship of the between-device differences with the BP levels. The differences of the two device readings were plotted against the corresponding averages of the two device readings for systolic and diastolic measurements separately (Bland and Altman graphs). Both figures show some extreme values beyond two SDs, but no discernible linear relationship could be ascertained. The Spearman rank correlation between the absolute difference and the means was 0.2 for systolic and –0.15 for diastolic. For systolic BP (Figure 5), the LOESS slightly deviates from the zero reference line at higher systolic values, indicating larger between-device



NOTE: BP is blood pressure in mm Hg.
SOURCE: CDC/NCHS, National Health and Nutrition Examination Survey, 2009–2010 blood pressure methodology study.

Figure 3. Scatter plot of Omron and mercury systolic blood pressure values overlaid by regression and unity lines

differences at higher BP values. In contrast with systolic BP, Figure 6 shows a curvilinear relationship for diastolic BP. Lower diastolic values were associated with larger between-device differences, but the differences decreased at higher diastolic values.

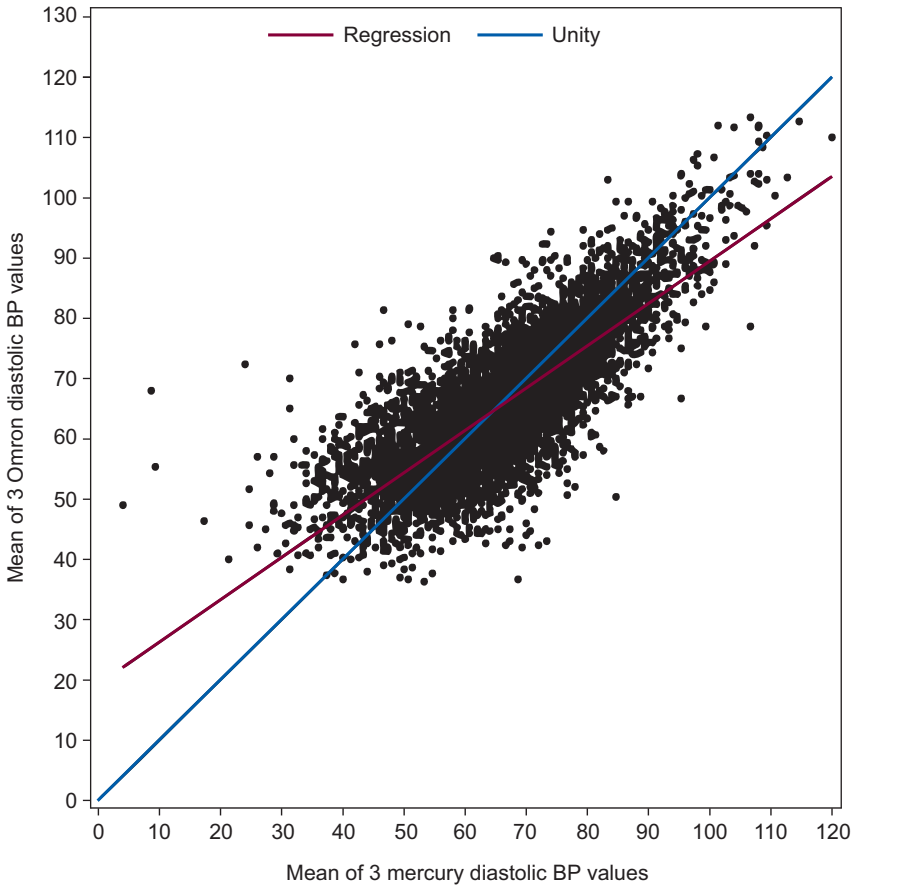
The means of the between-device differences are presented in Tables 3 and 4. Overall, the mean between-device differences (Omron minus mercury) for systolic was –1.6 mm Hg (SD = 6.8) and for diastolic was –0.6 mm Hg (SD = 7.8). The mean between-device differences by gender, age group, race and ethnicity, cuff size, BMI, and irregular heart rate were less than or about 2 mm Hg, except those for extra-large BP cuff (–3.1 mm Hg) and obese individuals (–2.6 mm Hg) for the systolic readings, and underweight individuals for the diastolic readings (–3.5). For the systolic readings, apart from persons classified as "other" and underweight, the mean between-device differences were statistically significant for all other subclassifications. As for diastolic readings, apart from females, persons aged 8–18 and 60 and over, non-Hispanic blacks, obese persons, and persons classified as having irregular heart rate, the mean between-device difference was statistically significant for all other subclassifications.

Comparing high BP classification agreement by device

Table 5 shows the between-device agreement for the classification of overall high BP (140/90 mm Hg or more) and stage II high BP (160/100 mm Hg or more), which was defined by BP taken in the MEC for persons aged 18 and over. The kappa agreements were 0.72 for both. For overall high BP, Omron correctly identified 70.28% of hypertensive individuals and 97.38% of normotensive individuals. For stage II high BP, Omron correctly identified 65.24% of stage II hypertensive individuals and 99.44% of stage II normotensive individuals. For both high BP categories, consistently more persons were correctly identified as nonhigh BP (higher specificity) than correctly identified as high BP (relatively lower sensitivity). The gold standard for the sensitivity and specificity analyses was mercury BP determinations. The percentages of high BP and stage II high BP based on the Omron measurements were lower than those based on the mercury measurements (12.87% compared with 15.15% for overall high BP; 2.98% compared with 3.75% for stage II high BP).

Discussion

The Omron and mercury device measurements were correlated (r = 0.92 for systolic BP and r = 0.79 for diastolic BP). Previous studies comparing the Omron device with mercury consistently showed lower readings of Omron for both systolic and diastolic BP (7–10). In our previous validation study comparing Omron with mercury, the results showed that the overall difference between the two devices was –1.62 mm Hg (SD = 6.14) for systolic and –1.64 mm Hg (SD = 6.63) for diastolic (10). The results of the current study show an overall similar underestimation using the Omron device for both systolic BP (–1.6 mm Hg, SD = 6.8) and diastolic BP (–0.6 mm Hg, SD = 7.8). Previous studies have also suggested that increased systolic values are associated with increased discrepancy between the Omron and mercury devices (7–10). The results from our earlier study suggested that both systolic and diastolic values were affected by increased average mercury BP values (12). The current



NOTE: BP is blood pressure in mm Hg.
SOURCE: CDC/NCHS, National Health and Nutrition Examination Survey, 2009–2010 blood pressure methodology study.

Figure 4. Scatter plot of Omron and mercury diastolic blood pressure values overlaid by regression and unity lines

study further supported these findings for systolic BP. However, for diastolic BP, the current study suggests that the relationship is not linear, but curvilinear, with a greater between-device difference at lower diastolic values (below 60 mm Hg) and little difference thereafter.

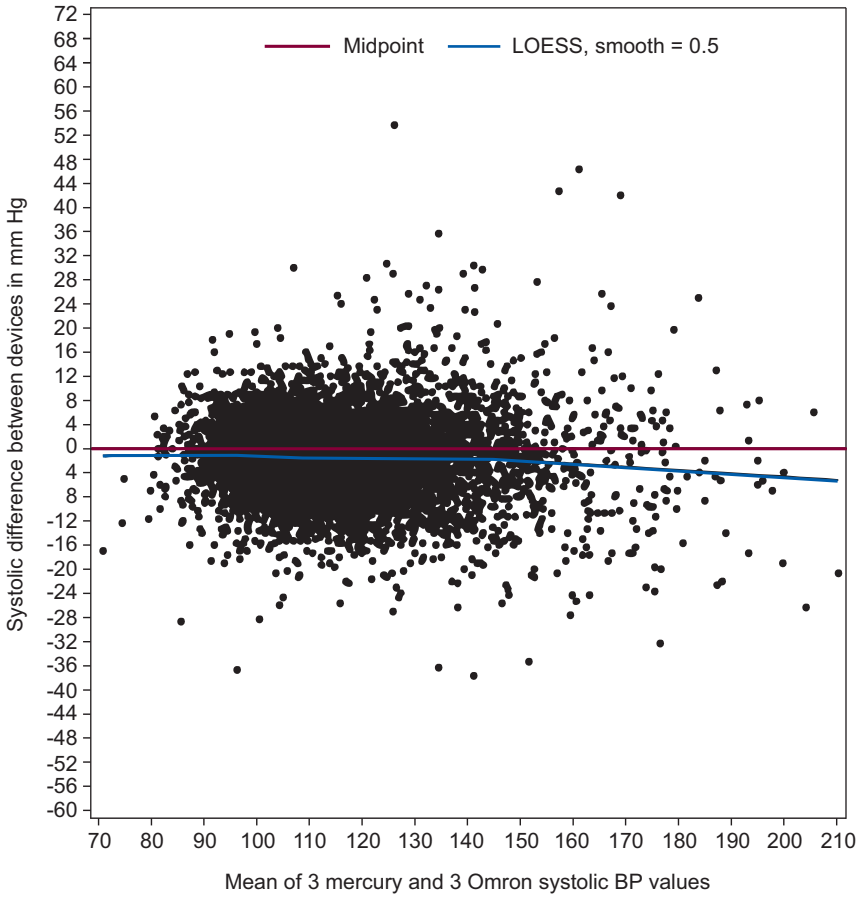
Although there was a beyond-chance agreement between the two device readings for high BP classification, the agreement was more likely to be for the absence of high BP (specificity) rather than the presence of high BP (sensitivity). Our kappa statistic for overall high BP classification in the current study was 0.72, a somewhat higher kappa than the kappa obtained in the previous study, which was 0.68. One reason for the low sensitivity and high specificity of Omron is that Omron reads lower than mercury in general. As a result, for persons who had BP at the borderline of high (140/90 mm Hg), Omron would read their blood pressure below 140/90 mm Hg and these subjects would be identified as normotensive, thus reducing the sensitivity. Similarly, for individuals who were measured as normotensive by mercury, it is more likely that they would be measured as normotensive by Omron, which leads to high specificity. To address this issue, we can adjust the Omron readings accordingly. Methods for adjustment include parametric methods, such as linear regression, demining regression, or a nonparametric method such as smoothing splines.

A total of 541 persons (8%) had a discrepancy in cuff-size selection when compared with midarm circumference values obtained during the body measurements exam in the MEC. Comparing the respective results after removing the individuals with discrepant cuff size suggested that removing these persons resulted in negligible changes (some slight changes after the decimal point) to no changes in values among respective BP determinations. Moreover, the prevalence of high systolic or high diastolic BP is nearly identical with and without exclusions. It also shows that at the high levels of BP, the mean values of between-device differences are nearly identical with and without exclusions.

As noted by Jones and others (19), oscillometric BP devices have problems obtaining BP readings in persons with arrhythmias. We attempted to have a proxy measure for atrial fibrillation by having the observers palpate the radial pulse—this crude measure suggested a 2 mm Hg mean between-device difference for systolic and less than 1 mm Hg mean difference for diastolic, although these results need to be interpreted cautiously. NHANES does not have complete information for atrial fibrillation. Available data in NHANES are one check box (irregular heartbeat) for the physician to check during the examination and lists of anti-arrhythmic medications in the household questionnaire. Based on NHANES 2007–2008 data only, 171 (1.8%) persons aged 8 years and over had an irregular pulse, and 35 persons (0.34%) in 2007–2008 were taking anti-arrhythmic medication.

Both the strengths and the limitations of the study can be related to its setting. The study was carried out during 2 survey years of NHANES, echoing the recommendations of Jones and colleagues (22) who advised that automatic BP devices should be assessed for accuracy under conditions of routine and frequent use. Unlike traditional validation studies, the BP observers were not continuously monitored for observer agreement within 4 or 5 mm Hg, nor did they repeat any measures with a difference greater than 4 or 5 mm Hg.

The study employed only standard mercury Baumanometer cuffs to perform



NOTES: BP is blood pressure in mm Hg. LOESS is local regression line.
SOURCE: CDC/NCHS, National Health and Nutrition Examination Survey, 2009–2010 blood pressure methodology study.

Figure 5. Scatter plot of between-device differences and mean of mercury and Omron systolic blood pressure values overlaid by local regression line

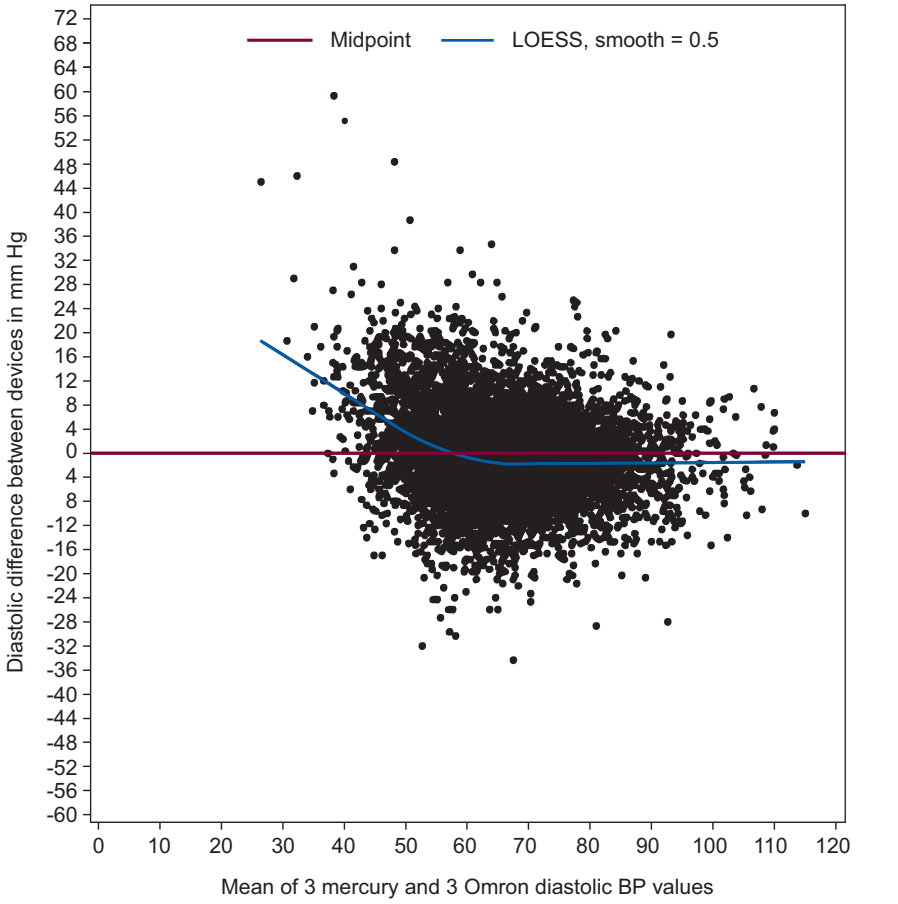
all BP measurements by both devices; our first comparison study followed the same method. The rationale for not changing the cuff was to control for the physiological reaction, which could be expressed if the cuff was changed in midmeasurement. Moreover, as previously described, the cuff sizes were selected according to predetermined arm circumference values. Lastly, the Omron device was calibrated against the mercury manometer using the four Bauman cuffs. In all of the 190 calibrations, the comparative readings were within ± 3 mm Hg of the mercury manometer readings adhering to the manufacturer's guidelines, which are more stringent that the AAMI guideline for device comparison (11).

Conclusions

Aside from observer effect on the BP readings, other factors may contribute to our findings, such as the fact that the physical basis for BP measurement using oscillometric devices differs from that of standard mercury measurements, and, therefore, systolic and diastolic pressure may not actually represent the same physiologic entity when measured using these two devices. The latter notwithstanding, with some exception the difference between devices was less than or about 2 mm Hg. The exceptions were systolic between-device differences for those using the extra-large BP cuff (–3.1 mm Hg) and for obese persons (–2.6 mm Hg). Another exception was the mean diastolic between-device difference for underweight individuals (–3.5 mm Hg). Between-device agreement for the frequency of high BP (140/90 mm Hg or more) and stage II high BP (160/100 mm Hg or more) was above chance (κ = 0.72 for both). Omron underestimated the high BP frequency by 2.28% and stage II high BP frequency by 0.77%. The Omron device was used under controlled circumstances, and results may not apply to use in other settings such as the physician office, clinic, or home. However, this study provided a great amount of experience using the Omron device in the usual NHANES environment. BP readings obtained with the Omron device provide lower estimates of high BP; therefore, future national prevalence estimates may need to be adjusted accordingly to take into account the between-device differences.

References

1. Messelbeck J, Sutherland L. Applying environmental product design to biomedical products research. Environ Health Perspect 108(Suppl 6):997–1002. 2000.
2. Zohoori N, Pulley L, Jones C, Senner J, Shoob H, Merritt RK. Conducting a statewide health examination survey: The Arkansas Cardiovascular Health Examination Survey (ARCHES). Prev Chronic Dis 8(3):A67. 2011.
3. Nieto FJ, Peppard PE, Engelman CD, McElroy JA, Galvao LW, Friedman EM, et al. The Survey of the Health of Wisconsin (SHOW), a novel infrastructure for population health research: Rationale and methods. BMC Public Health 10:785. 2010.
4. Oregon Health Study Group. Oregon health study (2009–2010). Available from: http://www.oregonhealthstudy.org/en/home.php.
5. Environmental Working Group, Health Care Without Harm. Protecting by degrees: What hospitals can do to reduce mercury pollution. 1999. Available from: http://www.bvsde.paho.org/bvsacd/cd37/15703.pdf.
6. National Institutes of Health. Mad as a hatter campaign for a mercury free NIH. Available from: http://www.nems.nih.gov/Pages/madhatter.aspx.
7. National Institutes of Health, National Heart, Lung, and Blood



NOTES: BP is blood pressure in mm Hg. LOESS is local regression line.
SOURCE: CDC/NCHS, National Health and Nutrition Examination Survey, 2009–2010 blood pressure methodology study.

Figure 6. Scatter plot of between-device differences and mean of mercury and Omron diastolic blood pressure values overlaid by local regression line

Institute. CARDIA: Coronary Artery Risk Development in Young Adults. Available from: http://www.cardia.dopm.uab.edu/.

8. White WB, Anwar YA. Evaluation of the overall efficacy of the Omron office digital blood pressure HEM–907 monitor in adults. Blood Press Monit 6(2):107–10. 2001.
9. Ombani S, Riva I, Giglio A, Caldara G, Groppelli A, Parati G. Validation of the Omron M5–I, R5–I and HEM–907 automated blood pressure monitors in elderly individuals according to the international protocol of the European Society of Hypertension. Blood Press Monit 12(4):233–42. 2007.
10. El Assaad MA, Topouchian JA, Darné BM, Asmar RG. Validation of the Omron HEM–907 device for blood pressure measurement. Blood Press Monit 7(4):237–41. 2002.
11. American National Standard, ANSI/AAMI SP10:2002. Manual, electronic, or automated sphygmomanometers. Arlington, VA: Association for the Advancement of Medical Instrumentation (AAMI). 2002.
12. Ostchega Y, Nwankwo T, Sorlie PD, Wolz M, Zipf G. Assessing the validity of the Omron HEM–907XL oscillometric blood pressure measurement device in a national survey environment. J Clin Hypertens (Greenwich) 12(1):22–8. 2010.
13. Grossardt BR, Graves JW, Gullerud RE, Bailey KR, Feldstein J. The occurrence of the alerting response is independent of the method of blood pressure measurement in hypertensive patients. Blood Press Monit 11(6):321–7. 2006.
14. National Center for Health Statistics. National Health and Nutrition Examination Survey analytic guidelines. 2004. Available from: http://www.cdc.gov/nchs/nhanes/nhanes2003-2004/analytical_guidelines.htm.
15. Omron Healthcare, Inc. Omron instruction manual: Model HEM–907XL. 2009. Available from: http://www.omronhealthcare.com/wp-content/uploads/hem-907xl_im.pdf.
16. National Center for Health Statistics. National Health and Nutrition Examination Survey (NHANES) questionnaire and exam protocol. http://www.cdc.gov/nchs/nhanes/nhanes_questionnaires.htm.
17. Ostchega Y, Prineas RJ, Paulose-Ram R, Grim CM, Willard C, Collins C. National Health and Nutrition Examination Survey 1999–2000: Effect of observer training and protocol standardization on reducing blood pressure measurement error. J Clin Epidemiol 56(8):768–74. 2003.
18. Bland JM, Altman DG. Measuring agreement in method comparison studies. StatMethods Med Res 8(2):135–60. 1999.
19. Cohen RA. An introduction to PROC LOESS for local regression. Paper 273. Cary, NC: SAS Institute, Inc. 1999. Available from: http://www.ats.ucla.edu/stat/sas/library/loesssugi.pdf.
20. O'Brien E, Atkins N, Stergiou G, Karpettas N, Parati G, Asmar R, et al. European Society of Hypertension international protocol revision 2010 for the validation of blood pressure measuring devices in adults. Blood Press Monit 15(1):23–38. 2010.
21. Pickering TG, Hall JE, Appel LJ, Falkner BE, Graves J, Hill MN, 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 45(1):142–61. 2005.
22. Jones DW, Appel LJ, Sheps SG, Roccella EJ, Lenfant C. Measuring blood pressure accurately: New and persistent challenges. JAMA 289(8):1027–30. 2003.

Table 1. Sample size for blood pressure methodology study: National Health and Nutrition Examination Survey, 2009–2010

Condition	Number of participants included	Number of participants excluded
Eligible to participate in blood pressure study[1]	7,410	. . .
Missing all blood pressure measurements (unit nonresponse)	. . .	343
Measurements obtained by excluded observer[2]	. . .	7
At least one valid blood pressure reading for both systolic and diastolic[3]	6,660	. . .
At least two valid blood pressure readings for both systolic and diastolic[3]	6,542	. . .
At least three valid blood pressure readings for both systolic and diastolic[3]	6,336	. . .
Had three valid systolic mercury and Omron readings	6,460	. . .
Had three valid diastolic mercury and Omron readings	6,338	. . .

. . . Category not applicable.
[1]Persons aged 8 years and over.
[2]Observers were excluded if too few observations were made.
[3]With both mercury and Omron devices.

NOTE: Mercury is the mercury sphygmomanometer device; Omron is the Omron HEM–907XL device.

SOURCE: CDC/NCHS, National Health and Nutrition Examination Survey, 2009–2010 blood pressure methodology study.

Table 2. Blood pressure values across selected percentiles, by device

Mean and percentile	Mercury	Omron	△ (Om – Me)
Systolic			
Mean of three blood pressure readings	117.3	115.7	–1.6
1%	86.7	86.3	–0.4
5%	94.0	93.0	–1
10%	98.0	96.7	–1.3
25%	105.3	103.7	–1.6
50%	113.0	113.0	–1.7
75%	124.7	124.7	–1.3
90%	138.7	138.7	–1.3
95%	148.0	148.0	–2
99%	172.0	172.0	–1.3
Diastolic			
Mean of three blood pressure readings	66.5	66.0	–0.5
1%	36.0	43.3	7.3
5%	46.0	49.3	3.3
10%	50.0	52.3	2.3
25%	58.7	58.0	–0.7
50%	66.7	65.0	–1.7
75%	74.7	73.0	–1.7
90%	82.0	80.7	–1.3
95%	87.3	86.0	–1.3
99%	98.0	96.3	–1.7

NOTE: Mercury (Me) is the mercury sphygmomanometer device; Omron (Om) is the Omron HEM–907XL device.

SOURCE: CDC/NCHS, National Health and Nutrition Examination Survey, 20092010 blood pressure methodology study.

Table 3. Means of between-device differences (Omron minus mercury) for systolic blood pressure values, by gender, age group, race and ethnicity, cuff size, BMI category, and irregular heart rate

Characteristic	Sample size	Mercury (SD)	Omron (SD)	Mean difference (SD)
Overall	6,460	117.3 (17.6)	115.7 (17.4)	† –1.6 (6.8)
Gender				
Male	3,231	119.4 (17.0)	118.3 (16.4)	† –1.1 (6.5)
Female	3229	115.1 (17.9)	113.0 (17.9)	† –2.1 (7.1)
Age group				
8–18 years	1,554	104.8 (10.1)	103.6 (9.8)	† –1.2 (5.7)
19–39 years	1,766	113.3 (12.2)	111.5 (12.0)	† –1.8 (5.9)
40–59 years	1,598	120.4 (16.2)	118.7 (16.1)	† –1.7 (6.7)
60 years and over	1,542	131.2 (19.4)	129.4 (19.2)	† –1.8 (8.5)
Race and ethnicity				
Hispanic	1,829	114.6 (17.2)	113.2 (17.2)	† –1.4 (6.5)
Non-Hispanic white	3,002	117.9 (17.1)	116.0 (16.6)	† –1.8 (6.8)
Non-Hispanic black	1,239	120.6 (18.8)	118.8 (18.7)	† –1.8 (7.1)
Other races	390	114.6 (17.0)	114.0 (17.5)	–0.6 (6.6)
Cuff size				
Child or small adult	343	100.8 (11.9)	99.4 (10.1)	† –1.3 (5.8)
Adult	2,089	112.8 (17.4)	111.7 (16.6)	† –1.2 (6.3)
Large adult	3,132	120.6 (16.9)	119.1 (17.0)	† –1.5 (6.9)
Extra-large adult	893	122.1 (16.2)	119.0 (17.0)	† –3.1 (7.6)
BMI category§				
Underweight	78	110.6 (21.7)	109.1 (18.4)	–1.5 (7.2)
Normal weight	1,415	117.6 (17.6)	116.2 (16.7)	† –1.5 (6.6)
Overweight	1,676	121.6 (17.2)	120.4 (17.2)	† –1.1 (7.0)
Obese	1,813	123.4 (16.9)	120.8 (17.5)	† –2.6 (7.4)
Heart rate				
Irregular heart rate	191	126.7 (19.0)	124.6 (18.5)	† –2.1 (6.9)

† $p < 0.05$.

§ Excludes persons under age 18.

NOTES: Mercury is the mercury sphygmomanometer device; Omron is the Omron HEM–907XL device. BMI is body mass index. SD is standard deviation.

SOURCE: CDC/NCHS, National Health and Nutrition Examination Survey, 2009–2010 blood pressure methodology study.

Table 4. Means of between-device differences (Omron minus mercury) for diastolic blood pressure values, by gender, age group, race and ethnicity, cuff size, BMI category, and irregular heart rate

Characteristic	Sample size	Mercury (SD)	Omron (SD)	Mean difference (SD)
Overall	6,338	66.5 (12.8)	66.0 (11.3)	† –0.6 (7.8)
Gender				
Male	3,166	67.8 (13.3)	66.8 (11.8)	† –0.9 (7.7)
Female	3,172	65.3 (12.0)	65.0 (10.6)	–0.2 (7.9)
Age group				
8–18 years	1,477	57.2 (10.8)	57.1 (7.5)	–0.2 (9.7)
19–39 years	1,759	68.0 (11.7)	67.2 (10.2)	† –0.8 (7.3)
40–59 years	1,592	73.6 (10.8)	72.1 (10.5)	† –1.5 (6.3)
60 years and over	1,510	66.4 (12.2)	66.7 (10.9)	0.3 (7.6)
Race and ethnicity				
Hispanic	1,783	64.8 (12.8)	64.7 (10.9)	–0.1 (8.0)
Non-Hispanic white	2,952	66.9 (12.2)	65.9 (10.8)	† –1.0 (7.4)
Non-Hispanic black	1,219	68.1 (13.7)	68.1 (12.3)	–0.0 (8.2)
Other races	384	67.0 (13.0)	65.7 (11.6)	† –1.3 (8.0)
Cuff size				
Child or small adult	306	54.8 (10.6)	53.4 (7.3)	† –1.4 (9.3)
Adult	2,047	63.3 (11.9)	61.8 (9.8)	† –1.5 (8.0)
Large adult	3,102	68.6 (12.2)	68.3 (10.6)	† –0.3 (7.3)
Extra-large adult	880	71.0 (13.1)	71.8 (11.3)	†0.8 (8.1)
BMI category§				
Underweight	78	65.8 (11.6)	62.4 (10.8)	† –3.5 (7.2)
Normal weight	1,403	67.4 (11.2)	65.5 (9.9)	† –1.9 (7.0)
Overweight	1,666	69.0 (11.6)	68.5 (10.4)	† –0.5 (6.8)
Obese	1,791	71.0 (12.6)	71.2 (11.0)	0.2 (7.4)
Heart rate				
Irregular heart rate	187	67.1 (13.2)	67.6 (12.3)	0.5 (7.7)

† $p < 0.05$.

§ Excludes persons under age 18.

NOTES: Mercury is the mercury sphygmomanometer device; Omron is the Omron HEM–907XL device. BMI is body mass index.

SOURCE: CDC/NCHS, National Health and Nutrition Examination Survey, 2009–2010 blood pressure methodology study.

Table 5. Classification of overall high blood pressure and stage II high blood pressure in persons aged 18 and over, by mercury and Omron devices

Device and agreement†		Overall high blood pressure		Stage II high blood pressure	
		Mercury		Mercury	
		Yes	No	Yes	No
Omron	Yes	532	111	122	27
	No	225	4,129	65	4,779
		Percent hypertensive			
Omron		12.87		2.98	
Mercury		15.15		3.75	
		Agreement statistics			
Sensitivity		70.28		65.24	
Specificity		97.38		99.44	
Kappa		0.72		0.72	

† Mercury blood pressure readings are gold standard.

NOTES: High blood pressure is systolic blood pressure of 140 mm Hg or higher, or diastolic blood pressure of 90 mm Hg or higher. Stage II high blood pressure is systolic blood pressure of 160 mm Hg or higher, or diastolic blood pressure of 100 mm Hg or higher. Mercury is the mercury sphygmomanometer device; Omron is the Omron HEM–907XL device.

SOURCE: CDC/NCHS, National Health and Nutrition Examination Survey, 2009–2010 blood pressure methodology study.

U.S. DEPARTMENT OF HEALTH & HUMAN SERVICES

Centers for Disease Control and Prevention
National Center for Health Statistics
3311 Toledo Road
Hyattsville, MD 20782

National Health Statistics Reports ■ Number 59 ■ October 5, 2012

Acknowledgments

This paper is dedicated to the memory of Dr. Lester R. (“Randy”) Curtin, a friend and a mentor—you are going to be missed. We would like to express our gratitude to Michele Chiappa, Technical Writer, Harris Corporation, for the review and initial editing of various drafts of this report. We also would like to thank the National Institutes of Health’s National Heart, Lung, and Blood Institute for its support and help in this project.

Suggested citation

Ostchega Y, Zhang G, Sorlie P, et al. Blood pressure randomized methodology study comparing automatic oscillometric and mercury sphygmomanometer devices: National Health and Nutrition Examination Survey, 2009–2010. National health statistics reports; no 59. Hyattsville, MD: National Center for Health Statistics. 2012.

For free e-mail updates on NCHS publication releases, subscribe online at: http://www.cdc.gov/nchs/govdelivery.htm.
For questions or general information about NCHS: Tel: 1–800–232–4636 • E-mail: cdcinfo@cdc.gov • Internet: http://www.cdc.gov/nchs
DHHS Publication No. (PHS) 2013–1250 • CS234654