White Paper Documenting the Use of a Blood Pressure Validation Study to Calibrate NYC HANES 2013-14 Blood Pressure Data

Introduction

The New York City Health and Nutrition Examination Survey (NYC HANES) was a population-based cross-sectional survey of NYC adults aged 20 years or older, modeled after the National Health and Nutrition Examination Survey (NHANES). The NYC Department of Health and Mental Hygiene (DOHMH) conducted the first NYC HANES in 2004, and NYC DOHMH and the City University of New York School of Public Health (SPH) conducted the second in 2013-14. (SPH staff, including NYC HANES co-PI, are currently affiliated with New York University Langone Health [NYU].) The goals of the survey were to monitor the health of NYC adults and to measure changes in health over time, including risk factors for chronic disease, like hypertension. Details about both surveys are published elsewhere.(1, 2)

In 2004, the physical examination was conducted in clinics, and blood pressure was measured using a mercury sphygmomanometer—the gold standard in measuring blood pressure. Due to secular trends of decreasing survey response rates nationally and locally, and to make taking the survey more convenient for participants, in 2013-2014 the NYC HANES exam was conducted at participants' homes. In order to avoid the hazard of a mercury spill in a participant's home, and to reduce the equipment weight carried by interviewers, Lifesource UA-789AC—an automatic blood pressure monitor—was used.

The Lifesource UA-789AC monitor was recommended by RTI International, the survey vendor for 2013-14 NYC HANES, as a validated instrument, and the device was tested for ease and durability by senior staff at the two institutions leading the study. After data collection, senior staff learned that the Lifesource vendor had found systolic and diastolic blood pressure readings from the UA-789AC to be on average 2.7 mm/Hg and 2.0 mm/Hg (respectively) lower than the mercury sphygmomanometer. However, these findings were not published by Lifesource. Though the Lifesource vendor had found lower readings compared to mercury, the 2013-14 NYC HANES blood pressure readings were substantially higher than the 2004 NYC HANES blood pressure readings. The percentage of 2013-14 participants who reported hypertension, however, was similar to the 2004 percentage. Furthermore, no similar increase in blood pressure had occurred nationwide between 2004 and 2014.(3) To ascertain whether some of the blood pressure differences between the two surveys were artifacts of the differences in measurement instrument, we sought to validate the Lifesource instrument against a standard mercury sphygmomanometer post NYC HANES data collection.

Between July 2017 and June 2018, NYU Langone Health and the NYC Department of Health and Mental Hygiene (DOHMH) conducted a blood pressure validation-calibration study. The purpose of the study was to compare blood pressure measurements using both devices—Lifesource UA-789AC and mercury sphygmomanometer—and to calibrate NYC HANES 2013-14 blood pressure data so that 2013-14 and 2004 results would be comparable.

Methods

Study population and recruitment

We performed sample size calculations prior to the study. With a suggested effective difference of 3.3 mmHg and a standard deviation of 6 for both systolic and diastolic blood pressure and a power of 85%, the goal was to recruit 183 participants. We aimed for an equal distribution of arm circumference groups, 22-29.9 cm, 30-37.9 cm, and 38-48 cm, and an equal distribution of high (≥130/80 mmHg) and normal (<130/80 mmHg) measured blood pressure values.

The study was conducted in two parts. The first part took place at an office of the NYC DOHMH between July and August 2017; employees had been asked to participate in the study. Volunteers were screened and only those 20 years or older, not pregnant, and those without a previous diagnosis of arrhythmia were eligible to participate. Two hundred and eighteen interviews were completed during this part of the study. The second part took place in East Harlem between April and June 2018 among a population known for high prevalence of uncontrolled hypertension. This population was a part of a longitudinal study conducted by NYU and NYC DOHMH and they consented to be contacted for future studies. We recruited this second group to ensure we had enough variation in blood pressure values, specifically, those in the higher range. Twenty interviews were completed during this part of the study.

In order to recruit participants at the NYC DOHMH, an email was sent to approximately 3,000 employees working at one of the Long Island City offices. The email briefly explained the study, including an incentive that would be provided, and included a link to a screening survey. Following completion of the survey, eligible participants were directed to schedule their own interviews using Acuity Scheduling software. Participants from East Harlem were phoned and invited to participate in the study. The study was approved by the Institutional Review Boards of NYC DOHMH and NYU School of Medicine. SurveyMonkey—an online cloud-based survey that can be used to create and securely administer surveys—was used for data collection.

Data collection

The Association for the Advancement in Medical Instruments (AAMI) 2013 guidelines for same arm sequential reading were followed when blood pressure was measured. Three interviewers were present in the room during the interview: 1) active interviewer, 2) passive interviewer, and 3) monitor. The active interviewer greeted participants; obtained informed consent; measured arm circumference, height, and weight; obtained the maximum inflation rate; placed the appropriate cuff; and controlled the deflation valve while reading blood pressure simultaneously with the passive interviewer. The passive interviewer obtained blood pressure measurements from the mercury device simultaneously with the active interviewer using a dual head stethoscope. The monitor collected demographic information from participants, operated the Lifesource UA-789AC monitor, and recorded its reading. To eliminate bias, passive and active interviewers were blinded from each other's mercury readings, and they both were blinded from the Lifesource UA-789AC readings. The order in which the devices were used during each interview was randomized, and interviewers were randomized to passive interviewer or active interviewer in each interview.(4)

Participants were seated in an upright position with their back resting against the chair and feet flat on the floor, and they sat quietly for 5 minutes before their blood pressure was measured. The quiet rest period was applied before using both the mercury sphygmomanometer and the Lifesource monitor. Blood pressure was measured three times with each device with 30 seconds intervals between each reading. Participants were compensated with a $10 gift card, if they finished all parts of the interview.

Cuff size

Mercury cuff size was determined based on arm circumference (AC) (mid-point between the acromion and olecranon processes) and following NHANES recommendations as follows:

- Child and small adult (extra-small) cuff for those with AC between 17 cm and 21.9 cm
- Adult (small) cuff for those with AC between 22 cm and 29.9 cm
- Large adult (medium) cuff for those with AC between 30 cm and 37.9 cm
- Thigh (large) cuff for those with AC between 38 cm and 47.9 cm

The first two categories were combined as small cuff size. Also, the large cuff was used for 4 participants who had AC≥48 cm.

Lifesource cuff size was determined based on AC and following device recommendations as follows:

- Small cuff for those with AC between 16.0 cm and 23.9 cm
- Medium cuff for those with AC between 24 cm and 35.9 cm
- Large cuff for those with AC between 36 cm and 44.9 cm
- Extra-large cuff for those with AC between 45 cm and 60 cm

Systolic and diastolic blood pressure

For each participant, readings from the active and passive interviewers were averaged. If readings from the two interviewers were different by more than 4 mm Hg, data were excluded. Data were excluded if only one interviewer was able to obtain a valid blood pressure reading. The final sample size was 213. Average systolic and diastolic blood pressures were calculated, excluding blood pressure measurements from the first attempt. If measurements from only one attempt were available, they were used as the average blood pressure. Excluding cases with only 1 blood pressure measurement from the sample in a sensitivity analysis did not affect findings or calibration, so they were included for power purposes.

Data analysis

Mean SBP and DBP was calculated for both mercury and Lifesource for the entire sample and according to mercury cuff size and BMI categories. The mean difference between the two devices (Lifesource-mercury) was compared to zero for systolic and diastolic blood pressure using paired t-test for the entire sample and according to mercury cuff size and BMI categories. The absolute between device difference was calculated (Lifesource-mercury) and rounded (down for <0.5 and up for ≥0.5) and then categorized based on both AAMI and the international European protocols for device comparison.(5) The correlation of blood pressure reading between the devices was calculated using Pearson correlation coefficient. Sensitivity and specificity as well as Kappa statistics were calculated to assess the agreement between the devices, where the mercury sphygmomanometer was considered the gold standard for systolic elevated blood pressure (≥130 mm Hg) and diastolic elevated blood pressure (≥80 mm Hg). Bland⁄Altman graphs—displaying between device difference on the Y axis and the means of the two device readings on the X axis—were used to test for systematic differences between the two devices.

To identify factors that predicted systematic differences between the devices, we constructed the following regression models:

- OLS regression: predicting mercury sphygmomanometer SBP and DBP by Lifesource SBP and DBP (respectively).
- GLM regression: predicting between device difference for SBP and DBP by Lifesource SBP and DBP (respectively).
- Robust regression: predicting between device difference for SBP and DBP by Lifesource SBP and DBP (respectively).

The best cross-over prediction model was selected based on p-values and R^2. All statistical analyses were done using SAS 9.4 for Windows (Cary, N.C).

Results

Descriptive

The majority of participants were 20 to 39 years old (51.9%) non-Latina white (40.9%) women (79.3%). For mercury cuff size, 1.4% required an extra-small cuff, 34.3% required small cuff (first two combined), 48.4% required a medium cuff, and 16.0% required a large cuff. 10.8% had SBP≥130 mm/Hg and 13.2% had DBP≥80 (table 1).

Table 1: study sample characteristics.

	N*	%
Total	213	---
Age groups		
20-39	110	51.6
40-59	79	37.1
60+	24	11.3
Gender		
Female	169	79.3
Male	43	20.2
Other	1	0.5
Race/Ethnicity		
Non-Latino white	87	40.9
Non-Latino black	45	21.1
Latino	38	17.8
Asian	34	16.0
Other	8	3.8
Refused	1	0.5
Mercury cuff size used		
Child and small adult (17-21.9 cm)	3	1.4
Adult (22.29.9 cm)	73	34.3
Large adult (30-37.9 cm)	103	48.4
Thigh (38-47.9 cm)	34	16.0
Lifesource cuff size used		
Small (16.23.9 cm)	6	2.8
Medium (24-35.9 cm)	154	72.3
Large (36-44.9 cm)	44	20.7
X-large (45-60 cm)	9	4.2
Arm circumference		
<22	3	1.4
22-29.9 cm	70	32.9
30-37.9 cm	106	49.8
38-47.9 cm	30	14.1
≥48 cm	4	1.9
Hypertension (mercury)		
Total (SBP≥130 or DBP≥80)	39	18.3
SBP≥130	23	10.8
DBP≥80	28	13.2
Hypertension (Lifesource)		
Total (SBP≥130 or DBP≥80)	40	18.8
SBP≥130	25	11.7
DBP≥80	35	16.4

Mean SBP and DBP measured by the Lifesource monitor were higher than those measured by the mercury sphygmomanometer; (111.9 vs 109.5, P<0.0001 for SBP, and 70.7 vs 68.7, P=0.0003 for DBP). When stratifying by mercury cuff size and BMI categories, there was no significant difference in SBP between mercury and Lifesource among those who required small cuff size and those with BMI<25. Lifesource measured higher SBP than mercury among those who required medium and large mercury cuff sizes and those with BMI≥25 (P<0.05 for all). Mean DBP measured by Lifesource was lower than that measured by mercury for those who required small mercury cuff size (P=0.003) and those with BMI<25 (P=0.002). DBP measured by Lifesource was higher than that measured by mercury among those who required medium and large mercury cuff size and those with BMI≥25 (P<0.05 for all) (table 2).

Table 2: average of systolic blood pressure (SBP) and diastolic blood pressure (DBP) using mercury sphygmomanometer and Lifesource UA-789 AC monitor and average between device difference according to cuff size and BMI categories.

		Mean (SD)			
	N	Mercury	Lifesource	Difference	P-value
SBP					
Total	213	109.5 (15.4)	111.9 (15.6)	2.4 (6.9)	<0.0001
Mercury cuff size					
Small	76	105.5 (13.9)	105.5 (14.1)	0.0 (5.7)	0.995
Medium	103	109.7 (15.8)	113.5 (14.9)	3.7 (6.9)	<0.0001
Large	34	117.8 (14.4)	121.3 (15.5)	3.5 (7.8)	0.012
BMI					
<25	77	102.6 (10.5)	103.2 (10.2)	0.6 (5.8)	0.330
25-29.9	60	107.8 (14.1)	110.9 (11.9)	3.1 (7.4)	0.002
≥30	76	117.9 (16.7)	121.5 (17.5)	3.6 (7.2)	<0.0001
DBP					
Total	213	68.7 (9.4)	70.7 (10.6)	1.9 (7.6)	0.0003
Mercury cuff size					
Small	76	68.6 (8.5)	66.9 (8.4)	-1.7 (4.7)	0.003
Medium	103	69.1 (9.6)	71.0 (10.3)	1.9 (6.3)	0.002
Large	34	68.0 (10.9)	77.8 (12.1)	9.8 (10.1)	<0.0001
BMI					
<25	77	67.1 (7.6)	65.6 (7.1)	-1.6 (4.4)	0.002
25-29.9	60	68.3 (10.0)	69.0 (9.4)	0.7 (6.4)	0.428
≥30	76	70.7 (10.3)	77.1 (11.2)	6.4 (8.7)	<0.0001

Figure 1 presents different categories for absolute between-device differences. For SBP, 26.3% had an absolute between-device difference within 2 mmHg, 59.2% had an absolute between-device difference within 5 mmHg, and 81.3% had an absolute between-device difference within 10 mmHg. For DBP, 29.1% had an absolute between-device difference within 2 mmHg, 63.4% had an absolute between-device difference within 5 mmHg, and 85.5% had an absolute between-device difference within 10 mmHg (figure 1).

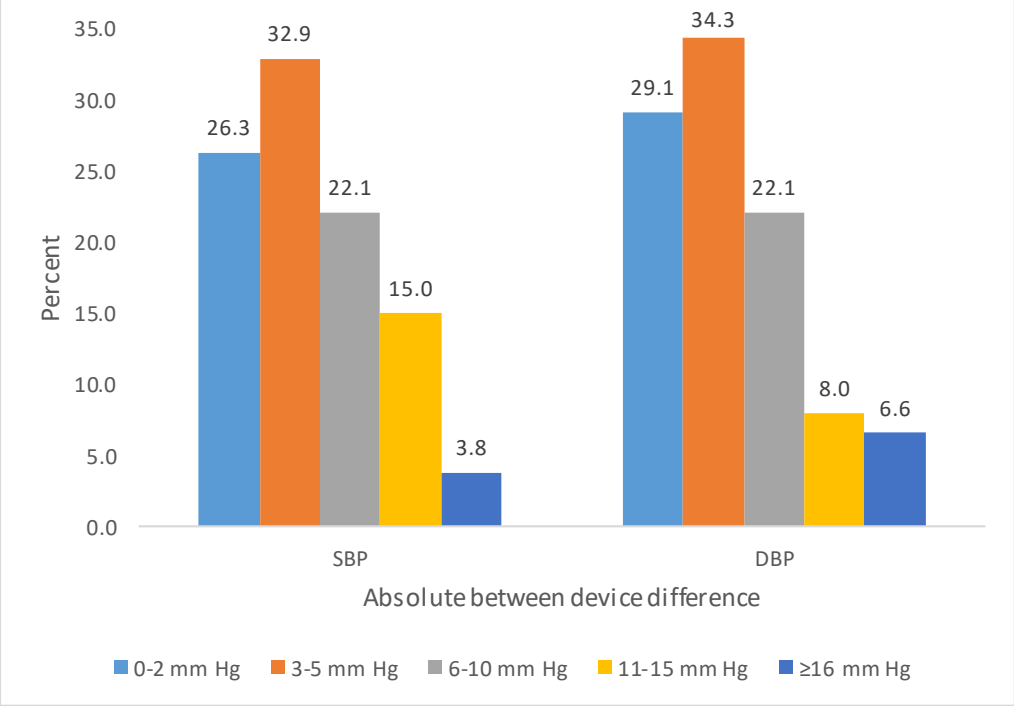

Figure 1: distribution of absolute between device difference (Lifesource-mercury) based on AAMI and international European protocol for device comparison

Kappa coefficient was 0.67 (95% CI: 0.54, 0.80), sensitivity was 74%, and specificity was 94% for total hypertension (BP≥130/80) classification. This translates to Lifesource identifying 74% of the true hypertensive subjects and 94% of those with normal blood pressure.

Figure 2 shows a linear correlation between the two devices for both SBP and DBP. Pearson correlation was 0.90 (P<0.001) for SBP and 0.70 (P<0.0001) for DBP.

A: Systolic blood pressure **B: Diastolic blood pressure**

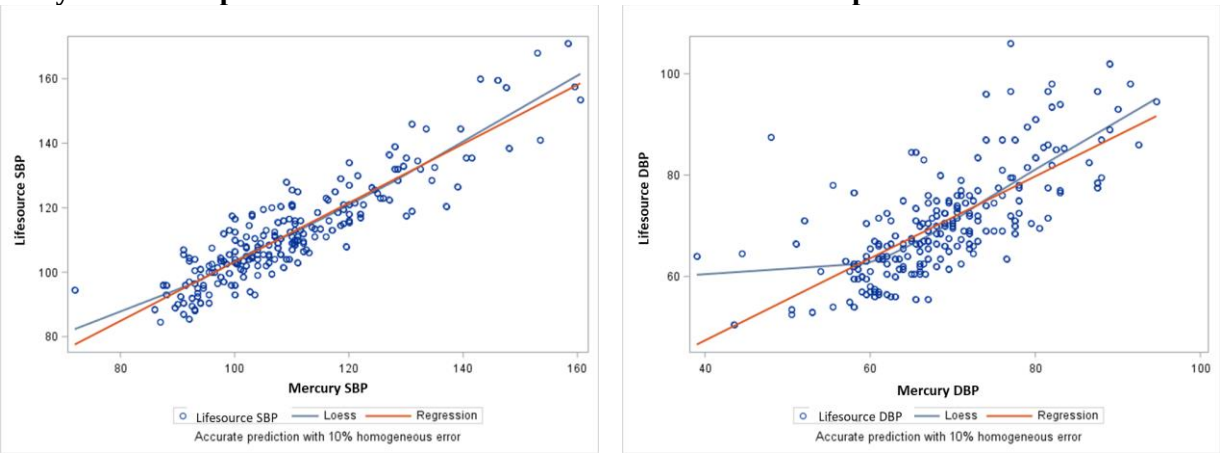

Figure 2: Scatter plot of Lifesource UA-789 AC monitor and mercury systolic and diastolic blood pressure.

The Bland-Altman plot did not show a linear relationship for between-device difference and the corresponding mean blood pressure from Lifesource and mercury, leading to the conclusion that there was no systematic difference between the two devices for either SBP or DBP (figure 3).

A: Systolic blood pressure **B: Diastolic blood pressure**

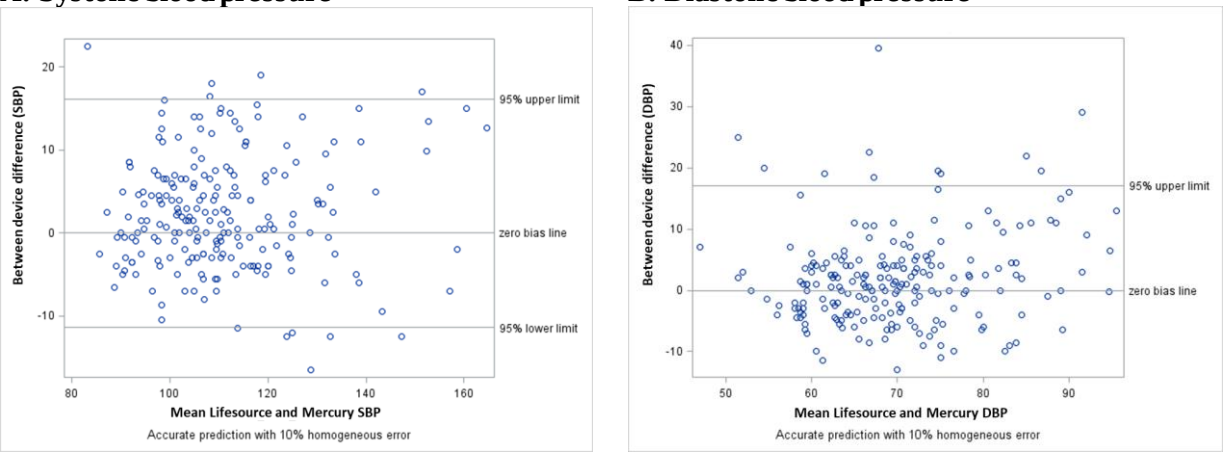

Figure 3: Blant-Altman graph, plotting between device difference and the corresponding mean of both devices for systolic and diastolic blood pressure

Regression models

Predicting mercury SBP, the best fit OLS and robust regression models did not control for any covariate and had R-squared of 0.81 and 0.08 (respectively). Among the GLM models tested to predict between-device difference in SBP by Lifesource UA-789AC, the best fit model included pulse pressure as a confounder and had an R-squared of 0.07. When predicting mercury DBP, among the OLS and GLM models test, the ones included gender and Lifesource cuff size as confounders were the best fit and had R-squared of 0.57 and 0.34 (respectively). While the best fit robust regression model included Lifesource

cuff size as a confounder and had an R-squared of 0.21.

Comparing R-squared and P-values across all models, the best prediction models were the OLS regression models for both SBP and DBP. NYC HANES 2013-14 blood pressure data were corrected using these models as follow:

$SBP_{mercury}=10.09+(0.89*SBP_{Lifesource})$ (R-squared 0.81)

$DBP_{mercury}=21.17+(0.68*DBP_{Lifesource})+(3.57*male)-(3.99*large\ cuff\ size)$ (R-squared 0.57)

Calibrating blood pressure data using the equations above lead to lowering the hypertension prevalence among NYC adults in 2013-14 (13.0% vs 21.9%).

Discussion

This study showed that, on average, the Lifesource blood pressure monitor measured SBP 2.4 mm Hg higher than the mercury sphygmomanometer and DBP 1.9 mm Hg higher than the mercury sphygmomanometer. In all, less than half of the sample had an absolute between-device difference greater than 5 mmHg and 5.2% of the sample were misclassified as hypertensive by Lifesource UA-789AC. In order to make blood pressure data from NYC HANES 2004 and 2013-14 comparable, we calibrated NYC HANES 2013-14 blood pressure data using OLS regression models. The calibration resulted in a significantly lower prevalence of hypertension among NYC adults in 2013-14.

The original Lifesource-led validation of Lifesource blood pressure monitor was not published in a peer reviewed journal. However, a comparable study conducted by the Centers for Disease Control and Prevention's National Center for Health Statistics compared blood pressure measurements between the mercury sphygmomanometer and another electronic automatic blood pressure monitor, the Omron HEM-907XL.(5) Authors of that study found the Omron HEM-907XL blood pressure monitor gave, on average, blood pressure readings that were 1.6 mmHg and 0.6 mmHg lower than the mercury sphygmomanometer for SBP and DBP (respectively). While agreement between the two devices was acceptable (72%), because the Omron HEM-907XL blood pressure monitor misclassified 2.3% of hypertensive individuals, the authors recommended calibrating data if comparison was to be made between blood pressure data when difference devices were used.

Strength of this included that we followed AAMI 2013 guidelines for validation in data collection to assure the accuracy of blood pressure measurements, as well as recommended analyses to compare the two devices. Although we met almost all of the AAMI guidelines, our data were limited for two groups: high blood pressure and large cuff size populations, and therefore our calibration of blood pressure data from these groups might not be robust. However, sensitivity analyses using different subsets of the data and different correction equations demonstrated similar results, suggesting that this limitation did not have a substantial effect.

In conclusion, we believe that the blood pressure differences between the two surveys were in part an artifact of using different blood pressure measurement devices. With the blood pressure validation study we were able to meet most AAMI validation guidelines, and successfully developed a model to correct the NYC HANES 2013-14 blood pressure data. These corrected data are publicly available, along with other study variables, at www.nychanes.org.

Funding source
This project was co-funded by NYU Langone Health faculty discretionary fund and the NYC Department of Health and Mental Hygiene.

Acknowledgements: We thank all the study participants, study staff Rodrigo Arce, Stephanie Rettig, Adeiyewunmi Osinubi, as well as the many DOHMH staff who made the study possible.

Co-Principal Investigators:
Rania Kanchi, MPH (Study leader)
Data Analyst
Department of Population Health
NYU School of Medicine

Sharon Perlman, MPH
Director of Special Projects
Division of Epidemiology
NYC Department of Health and Mental Hygiene

Lorna Thorpe, PhD
Professor
Director of Division of Epidemiology
Vice Chair of Strategy and Planning, Department of Population Health
NYU School of Medicine

Co-Investigators:
Shadi Chamany, MD, MPH
Science Director
Division of Prevention and Primary Care
NYC Department of Health and Mental Hygiene

Claudia Chernov, MPH
Data Analyst
Division of Epidemiology
NYC Department of Health and Mental Hygiene

Daichi Shimbo, MD
Associate Professor of Medicine
Director of the Translational Lab, Center of Behavioral Cardiovascular Health
Columbia University

Consultants and training personnel
Yechiam Ostchega, PhD, RN
Nurse consultant to the NHANES project
Center for Disease Control and Prevention

Tatiana Nwankwo, MS
Public Health Advisor
Centers for Disease Control and Prevention
National Center for Health and Statistics

References:

1. Thorpe LE, Greene C, Freeman A, Snell E, Rodriguez-Lopez JS, Frankel M, et al. Rationale, design and respondent characteristics of the 2013-2014 New York City Health and Nutrition Examination Survey (NYC HANES 2013-2014). Preventive medicine reports. 2015;2:580-5.
2. Thorpe LE, Gwynn RC, Mandel-Ricci J, Roberts S, Tsoi B, Berman L, et al. Study design and participation rates of the New York City Health and Nutrition Examination Survey, 2004. Preventing chronic disease. 2006;3(3):A94.
3. Fryar CD, Ostchega Y, Hales CM, Zhang G, Kruszon-Moran D. Hypertension prevalence and control among adults: United States, 2015–2016. NCHS data brief, no 289. Hyattsville, MD: National Center for Health Statistics. 2017.
4. Ni H, Wu C, Prineas R, Shea S, Liu K, Kronmal R, et al. Comparison of Dinamap PRO-100 and Mercury Sphygmomanometer Blood Pressure Measurements in a Population-Based Study*. American Journal of Hypertension. 2006;19(4):353-60.
5. Ostchega Y, Zhang G, Sorlie P, Hughes JP, Reed-Gillette DS, Nwankwo T, 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. 2012(59):1-15.