

Published in final edited form as:

Am J Hypertens. 2009 March ; 22(3): 263–266. doi:10.1038/ajh.2008.338.

Evaluating the Accuracy of an Aneroid Sphygmomanometer in a Clinical Trial Setting

Yong Ma¹, Marinella Temprosa¹, Sarah Fowler¹, Ronald J. Prineas², Maria G. Montez³, Janet Brown-Friday⁴, Mary L. Carrion-Petersen⁵, Tracy Whittington⁶, and The Diabetes Prevention Program Research Group

¹The George Washington University, Biostatistics center, Rockville, Maryland, USA

²Wake Forest University School of Medicine, Winston-Salem, North carolina, USA

³University of Texas health Science center, San antonio, Texas, USA

⁴Albert Einstein college of Medicine, Bronx, New York, USA

⁵University of california San Diego, San Diego, california, USA

⁶Johns hopkins University School of Medicine, Lutherville, Maryland, USA.

Abstract

BACKGROUND—The mercury sphygmomanometer, the “gold” standard for blood pressure measurements, has been gradually phased out in many institutions because of environmental concerns. Our on-going clinical trial compared the use of mercury vs. aneroid sphygmomanometers, before implementing a study-wide transition to the aneroid sphygmomanometer.

METHODS—The Diabetes Prevention Program Outcomes Study (DPPOS) studied the accuracy of the Welch allyn Tycos 767-Series Mobile aneroid sphygmomanometer from 20 March 2006 to 21 June 2006 at multiple clinic centers. We compared readings from 997 participants in 24 clinic centers using both mercury and aneroid sphygmomanometers.

RESULTS—The study found no statistically significant difference for systolic blood pressure (SBP) ($P > 0.05$) and a small but significantly ($P < 0.0001$) lower (0.8 mm hg) reading for diastolic blood pressure (DBP) using the aneroid sphygmomanometer. Regression analysis of aneroid vs. mercury showed regression lines ($Y = 4.8 + 0.96X$ for SBP, $Y = 3.1 + 0.95X$ for DBP) slightly but statistically significantly different from the line of equality ($P < 0.001$). Participants' age, sex, race/ethnicity, body mass index, blood pressure, and clinical center together explain about 8–10% of the variation of the difference between readings from the two sphygmomanometers. Based on the above result, on 1 august 2006, DPPOS clinics began the conversion from mercury to aneroid sphygmomanometers.

CONCLUSIONS—The Welch allyn Tycos 767-Series Mobile aneroid model 7670–04 tested in this validation study can be used to replace mercury model in clinical trials.

The mercury sphygmomanometer has historically served as the “gold” standard for measurement of blood pressure. Recently, because of environmental concerns about the disposal of mercury-contaminated medical waste and the risk of spills from mercury sphygmomanometers, the mercury sphygmomanometer is being phased out in clinical

settings. In 1998, the Environmental Protection Agency and American Hospital Association agreed to eliminate mercury containing waste from the health-care industry by 2005 (ref. 1). In response, many institutions began switching from mercury to aneroid sphygmomanometers.

The Diabetes Prevention Program was a controlled, multi-center clinical trial designed to compare the effect of intensive lifestyle change or medication on delay of diabetes onset in a prediabetic population.² The study ended in 2001 with the conclusion that intensive lifestyle change reduced diabetes risk by 58% and metformin medication by 31% over a period of 3 years.³ Participants continue to be followed in a long-term study: the Diabetes Prevention Program Outcomes Study (DPPOS)—designed to investigate longer-term complications of diabetes such as cardiovascular and microvascular outcomes (nephropathy, neuropathy, and diabetic retinopathy) over 10 years of follow-up.

Starting in October 2005, DPPOS clinics began to fall under state legislature mandates to eliminate mercury from health-care facilities. With the understanding that eventually all mercury sphygmomanometers will be replaced, the DPPOS study group investigated what effects such a conversion might have on the study's longitudinal blood pressure measurements. Although the published studies⁴⁻⁶ showed that aneroid devices could be accurate, several issues remained unanswered in relation to clinical trial studies. These prior aneroid validation studies were done by connecting aneroid sphygmomanometers to a reference unit and comparing the readings from the aneroid to the static pressure measurements set at fixed points. In real clinical trials, sources of variation in blood pressure measurements could include participant factors such as arm circumferences, technician factors such as measurement technique or interaction with participants. All of the above may affect the accuracy of aneroid models. In DPPOS 1–4 research staff collected measurements in each of 26 US clinics. To ensure the quality and consistency of our blood pressure data across clinics and over time, and to determine whether the shift to aneroid manometers would cause a systematic shift in measured blood pressure, the study group designed and implemented a study of its selected aneroid sphygmomanometer.

METHODS

The Diabetes Prevention Program was a randomized clinical trial involving participants at 27 centers who were at high risk for diabetes. Detailed methodology has been reported² and the protocol is available at <http://www.bsc.gwu.edu/dpp>. The institutional review board at each center approved the protocol, and all participants gave written informed consent before participation.

Mercury sphygmomanometer model

Before October 2005, DPPOS used the desk model gravity sphygmomanometer (300 mm Hg), manufactured by W.A. Baum, in Copiague, NY.

Aneroid sphygmomanometer model and calibration equipment

DPPOS selected the Welch Allyn Tycos 767-Series Mobile Aneroid Model 7670–04 as its aneroid sphygmomanometer.

The aneroid gauge consists of a metal bellows and a watch-like movement connected to the compression cuff. Variations of pressure within the system cause the bellows to expand and contract. Movement of the bellows rotates a gear that turns a pointer pivoted on bearings, across a calibrated dial.

The Netech DigiMano digital pressure and vacuum meter Part no. 200–2000IN was used for accurate calibration of the aneroid models. Calibration was done immediately before initiation of the comparative study.

Study procedure

We estimated that a sample size of 958 participants (31% of the DPPOS population) would provide 85% power to detect ≥ 2 mm Hg overall mean of the difference between the new sphygmomanometer and the standard, given an observed blood pressure standard deviation of 14.6 mm Hg.⁷ One clinic that had already converted to the aneroid model was excluded from this study. Each clinic was given a target number of participants to test and any participant who came to their annual or mid-year visit was eligible until the target was met.

BP technicians were routinely trained in using the mercury device. The study provided a training session for both devices with written protocols immediately before the start of this procedure. To reduce measurement bias, the order of sphygmomanometer type for measurement was randomly assigned. The same appropriate-sized cuffs, the same arm, and the same posture during measurement were used for both devices.

Over the period 20 March 2006 to 21 June 2006, the study conducted dual, sequential measurements in 1,062 participants. A resting (5 min) systolic and diastolic seated arm blood pressure was measured in the morning using both the new and the standard mercury devices. Because our participants were also undergoing testing of plasma glucose, they were studied in the fasting state. The dual BP measurements were separated by 30 s, and measurements were recorded on a data entry form. One clinic's data was excluded because readings from mercury and the aneroid were identical, caused by the technician's misunderstanding the protocol.

Statistical analysis

Each participant had four readings of systolic blood pressure (SBP) (two from each sphygmomanometer) and four reading of diastolic blood pressure (DBP). The average of the two repeated blood pressure readings was used for comparison of the two types of sphygmomanometer. Repeat readings from the mercury sphygmomanometer provided estimates of variation within participants and within technicians. Paired Student's *t*-test⁸ was used to test if the difference between the two types of sphygmomanometers is equal to zero. In addition, regression analysis⁹ was used to compare the two types. We used linear regression modeling to evaluate variation across clinical centers, demographic factors such as age, sex, and race/ethnicity, anthropometrics, body mass index, or blood pressure.

RESULTS

We analyzed data from 997 participants in 24 clinics who participated this study and had valid data. The average of the two readings from the mercury sphygmomanometers was 74.6 mm Hg (range 47–110) for DBP and 119.9 mm Hg (range 85–179) for SBP. The comparable aneroid sphygmomanometer means were: 73.8 mm Hg (range 44–115) for DBP and 119.8 mm Hg (78–179) for SBP. The mean difference (\pm s.d.) between the two sphygmomanometers (mercury-aneroid) was 0.8 ± 3.2 mm Hg ($P < 0.0001$) for the SBP and 0.1 ± 4.1 mm Hg for DBP ($P = 0.37$).

Figure 1 shows plots of readings from each of the two types of sphygmomanometers, separately for SBP and DBP. Regression analysis of aneroid BP vs. mercury BP revealed the following relationships:

$$\text{SBP (aneroid)} = 4.8 + 0.96 \text{ SBP (mercury)}$$

$$\text{DBP (aneroid)} = 3.1 + 0.95 \text{ DBP (mercury)}.$$

The slopes in both equations are significantly different from the line of equality (dotted line in Figure 1) ($P < 0.001$) and intercepts significantly different from zero ($P < 0.001$). Table 1 represents the individual predicted values from the aneroid model at various mercury values.

Table 2 presents the percent of instances in which the two devices differed by more than ± 3 and ± 5 mm Hg (the latter being the criterion endorsed by the British Hypertension Society¹⁰).

We explored within-participant variation by comparing back-to-back duplicate readings from the mercury device. The standard deviation of the difference between first and second readings was 2.8 for SBP and 3.7 for DBP. There was no significant difference between sequential systolic measurements within participant (mean at 0.2 mm Hg, $P = 0.20$). The second reading of DBP was slightly lower (0.3 mm Hg, $P < 0.01$) vs. the first (as might be expected from an accommodation effect). Table 2 shows the percent of the two mercury device measurements that were within ± 3 mm Hg and within ± 5 mm Hg. The variation within participants/technician pair explains most of the between-device variation.

Multiple linear regression analysis of the between-device difference in SBP adjusted for clinic, age, sex, race, body mass index, and SBP (measured from the mercury sphygmomanometer) showed that the between-device difference was positively associated with level of SBP (coefficient 0.05, 95% confidence interval (0.03, 0.07), $P < 0.0001$), and negatively with participant's age (coefficient -0.04 , 95% confidence interval $(-0.07, -0.01)$, $P = 0.002$) and varied by clinic center (P value < 0.0001 , least square means ranging from -4.3 to 4.4 , overall model $R^2 = 0.08$).

The same analysis of mercury vs. aneroid difference in DBP vs. clinic, age, sex, race, body mass index, and DBP (measured from the mercury sphygmomanometer) showed that the sphygmomanometer differences were positively associated with level of DBP (coefficient 0.06 95% confidence interval (0.04, 0.08), $P < 0.0001$) and varied by clinic center ($P < 0.0001$, with least square means range from -1.4 to 2.5 , overall model $R^2 = 0.10$).

Figure 2 plots the between-device difference against the readings from the mercury sphygmomanometer.

Based on the result that no clinically significant differences were observed between the two types of sphygmomanometers, the study changed to the aneroid sphygmomanometer on 1 August 2006.

DISCUSSION

This study demonstrated that aneroid sphygmomanometer Welch Allyn Tycos 767-Series model 7670-04 is comparable to mercury sphygmomanometer in a clinical trial setting without clinically significant systematic differences. We found no statistically significant difference between devices for SBP ($P = 0.37$) and a significant ($P < 0.001$) but small lower (0.8 mm Hg) DBP for the aneroid sphygmomanometer. Regression analysis detected statistically significant ($P < 0.001$) but small differences (see Figure 1).

Aneroid sphygmomanometers have more moving parts compared to mercury sphygmomanometers and are subject to fatigue. Bailey *et al.* reported that the aneroid sphygmomanometers (models not specified) selected were inaccurate in 80% of the units tested.⁴ In 1993, the Mayo Clinic, Rochester, Minnesota developed a standard protocol for regular calibration of aneroid devices. Canzanello *et al.* showed that aneroid

sphygmomanometers could provide accurate measurement of blood pressure when a proper maintenance protocol is followed,⁵ and that aneroid devices had a mean difference of 0.5 mm Hg lower than the reference levels, with 100% of aneroid readings falling within 4 mm Hg of the reference levels. Yarows and Qian⁶ also demonstrated that aneroid devices were accurate when calibrated annually. All studies connected the aneroid sphygmomanometer to a reference unit and then compared the aneroid readings to the static pressure measurements set at fixed points.

The present study compares measurement from aneroid and mercury sphygmomanometers in a clinical trial setting. On average, we found no clinically significant differences in between devices. For DBP, the reading from the aneroid sphygmomanometer was slightly lower. Regression analyses indicate that for both SBP and DBP, the readings from the aneroid will overestimate at the low end and underestimate at the high end, however, the differences are very small and clinically not significant. Because this study is to assess the accuracy of the aneroid model in a clinical trial setting, we did not exactly follow the validation protocol set by the Association for the Advancement of Medical Instrumentation or British Hypertension Society. However, when compared, this device is within the “pass” range set by Association for the Advancement of Medical Instrumentation (a mean ≤ 5 mm Hg and standard deviation ≤ 8 mm Hg) and grade A by British Hypertension Society (60% of the errors be within 5 mm Hg, 85% of the errors within 10 mm Hg, and 95% within 15 mm Hg).

This study is consistent with previous publications of aneroid accuracy based on directly connecting aneroid models with a standard calibrator.^{4,5} Variations among clinics reflect variation within technician, but only account for a small proportion of the overall variation. It appears that variation within participant/technician accounts for most of the variation in the blood pressure. We showed that for our selected aneroid device, accuracy could be maintained across 24 clinical centers with multiple research staff over with a wide range of blood pressure levels.

Acknowledgments

The Diabetes Prevention Program is supported by National Institutes of health/National Institute of Diabetes and Digestive and Kidney Diseases, the National Institute of child health and human Development, and the National Institute on aging; the Office of Research on Minority health and health Disparities, the Office of Women's health; the Indian health Service; the centers for Disease control and Prevention; the General clinical Research Program, the National center for Research Resources; the American Diabetes association; Bristol-Myers Squibb; Lipha Pharmaceuticals, Inc and Parke-Davis. We thank the thousands of volunteers in this program for their devotion to the goal of diabetes prevention. LifeScan Inc., health O Meter, hoechst Marion Roussel, Inc., Merck-Medco Managed care, Inc., Merck and co., Nike Sports Marketing, Slim Fast Foods co., and Quaker Oats co. donated materials, equipment, or medicines for concomitant conditions. McKesson BioServices corp., Matthews Media Group, Inc., and the Henry M. Jackson Foundation provided support services under subcontract with the coordinating center. A list of all staff involved with Diabetes Prevention Program is included on an online appendix.

References

1. Messelbeck J, Sutherland L. Applying environmental product design to biomedical products research. *Environ Health Perspect Suppl.* 2000; 108:S6.
2. The Diabetes Prevention Program Research Group. The Diabetes Prevention Program: design and methods for a clinical trial in the prevention of type 2 diabetes. *Diabetes Care.* 1999; 22:623–634. [PubMed: 10189543]
3. The Diabetes Prevention Program Research Group. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med.* 2002; 346:393–403. [PubMed: 11832527]
4. Bailey RH, Knaus VL, Bauer JH. Aneroid sphygmomanometers: an assessment of accuracy at a University Hospital and Clinics. *Arch Int Med.* 1991; 151:1409–1412. [PubMed: 2064493]

5. Canzanello VJ, Jensen PL, Schwartz GL. Are aneroid sphygmomanometers accurate in hospital and clinic settings? *Arch Intern Med.* 2001; 161:729–731. [PubMed: 11231706]
6. Yarows SA, Qian K. Accuracy of aneroid sphygmomanometers in clinical usage: University of Michigan experience. *Blood Press Monit.* 2001; 6:101–106. [PubMed: 11433131]
7. The DPP Research. Group. Hypertension, insulin, and proinsulin in participants with impaired glucose tolerance. *Hypertension.* 2002; 40:679–686. [PubMed: 12411462]
8. “Student” (W.S. Gosset). The probable error of a mean. *Biometrika.* 1908; 6:1–25.
9. Latman NS, Lanier R. Expressions of accuracy in the evaluation of biomedical instrumentation. *Biomed Instrum Technol.* 1998; 32:282–288. [PubMed: 9619256]
10. O'Brien E, Petrie J, Littler WA, de Swiet M, Padfield PL, Altman D, et al. The British Hypertension Society protocol for the evaluation of blood pressure measuring devices. *J Hypertens.* 1993; 11(suppl 2):S43–S63.

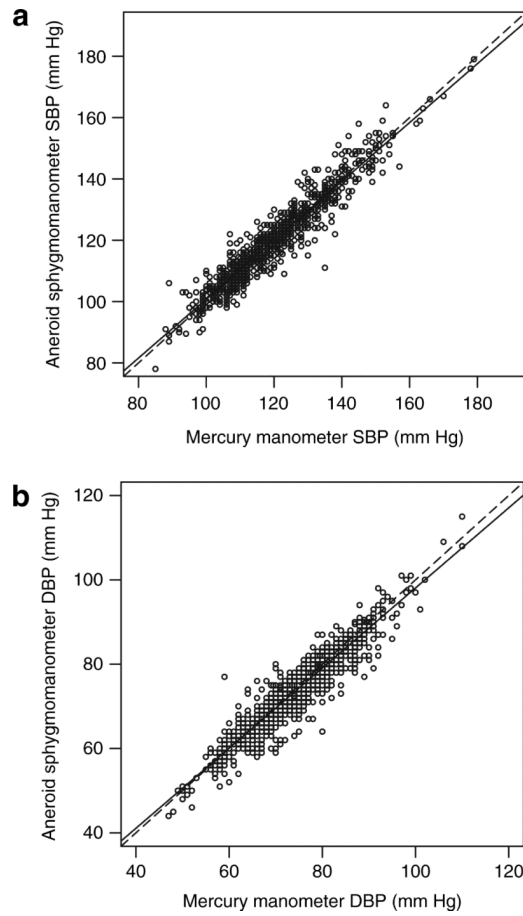


Figure 1. Comparison of the readings from the two sphygmomanometers. The dotted lines represent line of equality and the solid lines represent the regression of Y on X .

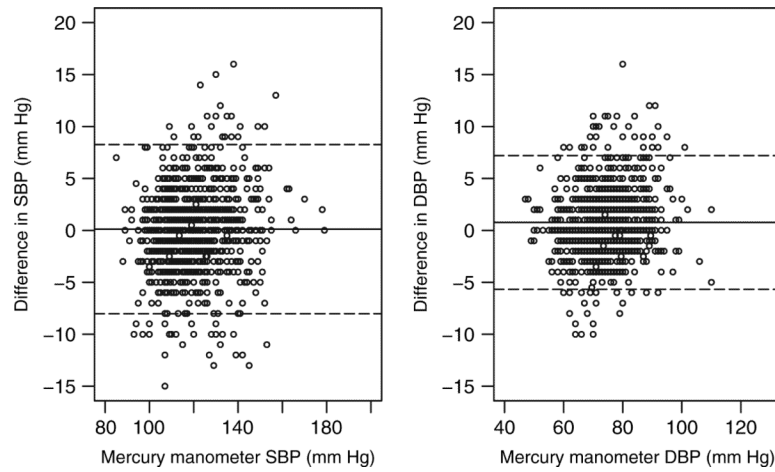


Figure 2. Plot of the difference between the two sphygmomanometers against the mercury readings. The middle solid line represents the mean of the values on the y-axis and the two dashed lines represent the 2 s.d. of the values on the y-axis.

Table 1

Predicted aneroid readings from regression analysis

Mercury value (mm Hg)	Predicted individual aneroid value (95% CI*) (mm Hg)
SBP	80
	82 (74, 89)
	100
	101 (93, 109)
	120
	120 (112, 128)
	140
	139 (131, 147)
	160
	158 (150, 166)
	180
	177 (169, 185)
DBP	40
	41 (35, 47)
	60
	60 (54, 66)
	80
	79 (73, 85)
	100
	98 (92, 104)
	120
	117 (111,123)

CI, confidence interval; DBP, diastolic blood pressure; SBP, systolic blood pressure.

Table 2Percent of the data with the absolute difference ≤ 3 and ≤ 5 mm Hg

	Total, N	N (%) (≤ 3 mm Hg)	N (%) (≤ 5 mm Hg)
Difference by type of sphygmomanometer (mercury-aneroid)*			
SBP (mm Hg)	997	683 (68.5%)	853 (85.6%)
DBP (mm Hg)	997	778 (78.0%)	901 (90.4%)
Difference in mercury measurement (1st-2nd)			
SBP (mm Hg)	997	641 (64.3%)	854 (85.7%)
DBP (mm Hg)	997	775 (77.7%)	926 (92.9%)

* Average of the two mercury readings minus the average of two aneroid readings.