PRAGMATIC VERSUS STANDARDISED BP MEASUREMENT: AN ANALYSIS OF BP MEASUREMENT IN A PRIMARY CARE HOSPITAL IN SWAZILAND

Ganizani Mlawanda[a, b], Michael Pather[a], Srini Govender[a,c]

Author affiliations

a) University of Stellenbosch, Faculty of Medicine, Department of Interdisciplinary Health Sciences, Division Family Medicine and Primary Care, Tygerberg, Cape Town, South Africa.

b) Royal Swaziland Sugar Corporation (RSSC) Medical Services, Mhlume, Swaziland.

c) Khayelitsha Community Health Centre, Khayelitsha, Cape Town, South Africa.

Sources of support and disclosure of funding

None.

Competing interests by author

None

Research Ethics committee approval

Ethics approval was granted by University of Stellenbosch Human Research Ethics Committee (HREC) N10/11/394 on 13th of May 2011 as well as institutional ethical approval.

Author's contribution

Dr Ganizani Mlawanda conceived the study, formulated the study design, data collection, statistical analysis and manuscript design. Statistical analysis was further verified by University of Stellenbosch statistician, Professor Justin Harvey. Dr Michael Pather and Dr Srini Govender were supervisors and gave authorization for Ethics and final submission. All authors read and approved the final manuscript.

DECLARATION

I, the undersigned, hereby declare that the work contained in this assignment is my original work and that I have not previously submitted it, in its entirety or in part, at any university for a degree. I also declare that approval for the study was obtained from the Human Research Ethics Committee of Stellenbosch University (Reference number N10/11/394).

_____ _____

Dr Ganizani Mlawanda Date

ABSTRACT

Background: Measurement of blood pressure (BP) is done poorly due to both human and instrument errors. The standard protocol for measurement is often not followed by healthcare workers.

Objectives: There were three main objectives: firstly to assess the difference between BP recorded in a pragmatic way and that recorded using standard BP measurement guidelines; secondly to assess difference between BP measurements done by wrist sphygmomanometer compared to mercury sphygmomanometer; and finally to assess if the differences affect decision to start or adjust hypertension treatment.

Setting: RSSC Mhlume hospital, Swaziland

Study design: cross sectional study

Study Method: Following consent, BP was assessed in a pragmatic way by nurse practitioner who made treatment decisions. Thereafter, patients had BP re-assessed using standard BP protocol by mercury (gold standard) and wrist sphygmomanometer. In addition demographic and clinical data was collected.

Results: The prevalence of hypertension was 25%. The mean systolic BP was 143 mmHg for pragmatic BP, 133 mmHg for standard BP using mercury sphygmomanometer and 140 mmHg for standard BP assessed using wrist device. The mean diastolic BP was 90 mmHg, 87 mmHg and 91 mmHg for pragmatic, standard mercury and wrist respectively. Pearson and intra-class correlation coefficients were similar for both systolic and diastolic BP and for all BP measurement pairs which were being compared. Bland Altman analyses showed that pragmatic and standard BP measurement were different and could not be used interchangeably. Standard mercury and wrist based methods were not clinically interchangeable. Treatment decisions between those based on pragmatic BP and standard BP agreed in 83.3% of cases; 16.7% of participants had their treatment outcomes misclassified. Twenty-five percent of patients were erroneously started on anti-hypertensive therapy based on pragmatic BP.

Conclusion: There is a difference between pragmatic and standard BP measurements which affect decisions not to start treatment and decision to start treatment but not treatment alteration decision for those already on treatment. There are also marked differences between wrist and standard mercury based BP devices. Clinicians need to revert to basic good practice and measure BP more accurately to avoid unnecessary additional costs and morbidity associated with incorrect treatment due to disease misclassification. Contrary to existing research, wrist devices need to be used with caution.

Key words: BP measurement methods; Pragmatic BP; Standard BP; wrist sphygmomanometer;

INTRODUCTION

Hypertension is a powerful, consistent, and independent risk factor for cardiovascular disease, stroke and renal disease. [1] Diagnosis of hypertension is based on measurement of blood pressure (BP). Obtaining accurate BP readings was noted to be a challenge faced by health professionals at all levels. [2] Numerous surveys have shown that physicians and other health care providers rarely follow established guidelines for BP measurement. [3] A look into variations between pragmatic ("real-life") and standardised (as per protocol) BP measurement will be useful in improving chronic disease management and ensuring the limited resources in our set up are used effectively. Technology has brought in various BP measuring devices, a common one in primary care being the wrist sphygmomanometer as opposed to the "gold standard" but environmentally unfriendly mercury sphygmomanometer. How does BP measurement from wrist device compare with the "gold standard"?

LITERATURE REVIEW

Hypertension is a global health burden affecting developed and developing nations. [4] The prevalence of high BP increases dramatically with age, such that the lifetime risk of high BP approaches 100%. [5] Extensive data have shown beyond doubt the benefit of controlling hypertension. [6]

Control of BP begins with accurate measurement that leads to appropriate diagnosis, assessment of cardiovascular risk and treatment decisions. [1-6] The target BP for patients using antihypertensive treatment has been lowered for those with diabetes or renal disease. [1] Thus, it has become increasingly important to be able to detect small differences in BP. Whilst BP measurement is a vital clinical skill, it is poorly performed by all health care professional categories. [4] In general, there are three sources of error in the indirect measurement of BP: *(a)* observer bias, *(b)* faulty equipment, and *(c)* failure to standardize the techniques of measurement by clinicians. [7]

Because of its accuracy and reliability, the mercury sphygmomanometer is generally regarded as the gold standard against which all other devices for blood pressure measurement should be compared. [5] Due to environmental issues, there has been increasing pressure to remove medical devices containing mercury from clinical areas, which is leading to the demise of the mercury sphygmomanometer, and as a result, automated BP devices have been embraced by clinicians for their convenience and ease of use. [8]

Rose et al suggested that the observer was the most critical component of accurate blood pressure measurement. [9] The British Hypertension Society declared that only an observer who is aware of the factors

that lead to false readings should measure BP because wrong readings obtained through failure to use the proper technique often lead to the wrong diagnosis, which may result in unnecessary or inappropriate treatment and follow up. [10]

Compared to physicians, trained nurses overestimated, rather than underestimated, blood pressure, although systolic blood pressure underestimation was extremely prominent in participants with moderate to severe hypertension. [11] Systolic blood pressure underestimation of >5 mmHg was as high as 57.5% by trained nurses using the traditional device versus 33.8% by the automatic device, indicating that nurses tended to underestimate blood pressure in participants with more severe hypertension. [11] The BP done by nurses was found to be consistently higher than that recorded by doctors. [11] McKay et al noted that few physicians have patients rest for 5 minutes before blood pressure measurement as recommended and as a consequence BP done by doctors was consistently high due to white coat effect. [12] Contrary to the recommended five minutes of rest, it appears that ten minutes rest before clinic BP evaluation could improve further the precision and accuracy of the measurement and implies that the optimal time at rest before clinic blood pressure (BP) measurement is still undefined. [13]

Clinicians should also be aware that BP in human beings is affected by multiple physiologic stimuli such as respiration, temperature, body posture, emotional or physical stress, meals, alcohol, or caffeine and smoking and hence should take these factors into consideration during measurement of BP. [14] For some patients, BP measurements taken in a doctor's office may not correctly characterize their typical BP. In up to 25% of patients, the office measurement is higher than their typical BP, a phenomenon known as white-coat hypertension. [14]

From the literature reviewed, it is clear that BP measurement is subject to errors. Thus there are still some social and scientific questions which need clarity and further research especially in resource limited settings. Literature review concluded that with proper measurement technique, machine variation between the gold standard mercury sphygmomanometer and the wrist is minimal. [3,7,10] In addition there are problems associated with pragmatic nature BP measurement and other observer related errors. [1-10]

Nearly all the articles found on literature review are from developed countries with a good patient to health-worker ratio. In a developing country setting, where the patient to health-worker ratio is low and resources limited, the potential for BP measurement errors may be worse. One obvious question was on assessment of the reliability of BP measurement methods looking at both sphygmomanometer and observer differences in resource-limited settings. In so doing, such research will further enlighten health workers about the

trustworthiness of BP readings and ensure that health workers are treating BP optimally. Problems related to over or under treatment may be serious and if identified early could reduce unnecessary morbidity and mortality. Most of the prior studies have mainly focused on sphygmomanometer related differences.

STUDY RATIONALE AND MOTIVATION

An analysis of variations between pragmatic or "real life" and standard BP measurement based on the "gold standard" would be useful in improving chronic disease management and ensuring effective use of already strained resources in primary care. A study of this nature may have an impact on increasing awareness of human induced variation in BP measurement and its impact on therapeutic decisions. Hence it may motivate clinicians to follow protocol. In the long run, it may have some economic advantages in saving cost of drugs erroneously prescribed to those who, if BP had been recorded properly, would not need treatment.

AIMS AND OBJECTIVES

Research Question

What is the difference between pragmatic and standard BP measurement in primary care?

Aims

1) To ascertain variations between standard and pragmatic BP measurements and comparison of wrist BP and mercury sphygmomanometer based BP.
2) To assess the impact of any differences on treatment decision

Objectives

1) To quantify the existence of any differences between BP recorded in a pragmatic way and that recorded using standard BP measurement protocols.
2) To quantify any discrepancy between BP measurements done by wrist sphygmomanometer when compared to mercury sphygmomanometers.
3) To assess if the differences in BP measurement have impact on treatment decisions: whether not to treat, to start antihypertensive treatment or adjust hypertension treatment.

METHODS

Study Design: This study was a cross sectional (observational) study.

Study Setting: This study was done at RSSC Mhlume hospital targeting outpatient. RSSC Mhlume hospital is a rural primary care facility in eastern part of Swaziland.

Study Population: The study population were adult (> 18years) patients, with or without hypertension, who attended primary care at the RSSC hospital during the study period June 2011 to December 2011 and gave consent to participate in the study.

Sample size and sampling method: Every forth patient who had attended the outpatient clinic was eligible for selection. A sample size of 60 was used: this based on statistical calculations and sample size from similar studies. [15] Statistically, subjects with two observations per subject achieves 80% power to detect an intra-class correlation difference of 0.15 using an F-test with a significance level of 0.05. In a similar study of agreement, Bland and Altman recommend a sample size of 30 as a "good sample" and 60 as "excellent" as it gives a 95% Confidence Interval about +/-0.34s, where s is the standard deviation of the differences between measurements by the two methods. [15]

Data collection and measurement method: Informed consent was obtained from eligible patients. Participant had BP assessed in a pragmatic way by nurse practitioners who would give their therapeutic decision based on their readings. Participants had BP reassessed according to the standard protocol using mercury sphygmomanometer and wrist sphygmomanometer alternately. To reduce bias, the order of measurement for pragmatic or standard BP measurements was alternated for successive patients. Finally demographic and relevant clinical data was collected into a "Data Collection" form, which was subsequently entered into an MS Excel for analysis.

How bias was minimized: To improve internal validity, the potential biases where handled as follows:

Selection bias- to reduce selection bias a systematic random sample (every forth patient) was used.

Measurement bias-this could occur with measurement, recording, management or analysis of the data. Notable were Hawthorne effect (nurses could change their BP measurement routine because they are aware of the investigation going) and observer diagnostic suspicion bias. These were reduced by blinding the nurse researcher of results from nurse practitioners; nurse practitioners were blinded of the ongoing study. Use of validated, standardized and calibrated sphygmomanometers reduced instrument variation. Batteries for the wrist devices were replaced regularly. To reduce subject physiologic variation, and the known

regression to mean with repeated BP measurement phenomenon, [16] the standard BP was measured within a few minutes before or after the pragmatic BP.

Confounding- Time between performing the BP measurements was an important confounder. Blood pressure tends to come down with time - regression to the mean. The time between pragmatic and standard BP assessment was kept at a minimal to reduce the possibility of confounding bias. Previous studies indicate that a time lag of less than ten minutes does not affect the BP result significantly. [13]

Data/Statistical Analysis: MS Excel was used to capture the data and STATISTICA version 9 (StatSoft Inc. (2009) STATISTICA (data analysis software system), www.statsoft.com.) was used to analyze the data. The statistical analysis comprised of descriptive and analytical statistics. For descriptive statistics, summary statistics were used to describe the variables. Medians or means were used as the measures of central location for ordinal and continuous responses and standard deviations and quartiles as indicators of spread. Wilcoxon sign rank test was used to assess differences between means of BP. For analytical statistics, simple logistic regression; Pearson correlation, intra-class correlation coefficient (ICC) and Kappa were used appropriately. Standard reference scales were used for Pearson, ICC and Kappa. Bland and Altman (BA) method of analysis of agreement was used for further assessment of agreement. Reference ranges for comparison of BA analysis were: within 10 mmHg for diastolic BP and within 20 mmHg for systolic BP because these are known ranges for hypertension severity grading. [4,5,6] Throughout the analysis, a p-value of $p < 0.05$ represented statistical significance in hypothesis testing and 95% confidence intervals were used to describe the estimation of unknown parameters.

RESULTS

A total of 60 out-patients consented to participate in the study. Thirty two were males. The mean age was 42.6 years. The mean weight was 77.8 kg and the mean height was 1.6 metres. The prevalence of hypertension was 25%. Twenty eight percent of the participants had co morbid diseases. The mean systolic BP was 143 mmHg for pragmatic BP, 133 mmHg for standard BP using mercury sphygmomanometer and 140 mmHg for standard BP assessed using wrist device. The mean diastolic BPs was 90 mmHg, 87 mmHg and 91 mmHg for pragmatic, standard mercury and wrist respectively. It took an average of 4.2 minutes between pragmatic and standard BP measurement. Table 1 below summarizes the findings. Three participants reported either a full bladder or had eaten within 30 minutes before BP assessment, five had exercised, one had smoked and taken coffee and seven reported some degree of psychological stress.

Table 1: Demographic and clinical characteristics of participants

Demographic and clinical characteristics					
Gender	Males	32	Females	28	
Mean Age in years (standard deviation):	43 (14.2)				
Mean Weight in kg (standard deviation):	78 (19.4)				
Mean Height in cm (standard deviation):	164 (8.5)				
Mean BMI	29				
Hypertensive patients	25%				
Co morbid conditions	28%				
MUAC in cm	32				
Treatment plan based on pragmatic BP - *no treatment:*32 (53%) *Treat*:18 (30%) *Change treatment*:10 (17%)					
Treatment plan based on standard BP - *no treatment*: 41 (68%) *Treat*:11 (18%) *Change treatment*:8 (13%)					
Systolic BP results	***Observation***	***Mean***	***25th Centile***	***50th Centile***	***75th Centile***
Pragmatic BP	60	143	120	140	163
Mercury Standard BP	60	133	110	130	151
Wrist Standard BP	60	140	123	138	155
Diastolic BP results	***Observation***	***Mean***	***25th Centile***	***50th Centile***	***75th Centile***
Pragmatic BP	60	90	73	90	105
Mercury Standard BP	60	87	75	85	102
Wrist Standard BP	60	91	77	88	106
Mean time between pragmatic and standard BP in minutes	4				

There were some differences in Systolic and Diastolic BP between pragmatic, standard mercury and wrist device BP. Because an assessment of Gaussian distribution for the difference between the BPs were non-Gaussian, a distribution free test, Wilcoxon sign rank test, was subsequently used to assess the association. The Wilcoxon sign rank test tests the equality of matched pairs of observations, the null hypothesis being that both distributions are the same. Table 2, summarizes the findings.

Table 2: Assessment of difference in BP between pragmatic, standard mercury and standard wrist BP

BP methods in comparison	**2-sided p-value(*interpretation)**	**2-sided p-value (*interpretation)**
	Systolic BP	***Diastolic BP***
Pragmatic BP/Standard Mercury BP	0.00 (difference)	0.02 (difference)
Pragmatic BP/Standard Wrist BP	0.17 (no difference)	0.44 (no difference)
Standard Mercury/Standard Wrist BP	0.00 (difference)	0.00 (difference)

***Interpretation based on p-value = 0.05.** Null hypothesis-H_0: null hypothesis states that, median difference of BP between any two given BPs (pragmatic, standard mercury and standard wrist BPs) is equal to zero. Reject H_0 if $p < 0.05$ and conclude there is a difference. Fail to reject H_0 if $p > 0.05$ and conclude that there is no difference.*

Thus, there was a statistically significant difference in systolic and diastolic BP between standard mercury BP and both pragmatic BP and wrist BP. On the contrary, there was no statistically significant difference for both systolic and diastolic BP between pragmatic and standard wrist BP.

Analytical results were as follows: the Pearson correlation coefficient (***r***) was the same, 0.9, for systolic and diastolic BP for all BP methods which were being compared, that corresponded to "good association" between pairs being compared. The Intra-class correlation coefficient (model 2) was consistent with "almost perfect agreement" for all methods compared. Thus ***r*** and ICC could not differentiate further the level of agreement between the methods in study. Adjustment for confounding was done: neither psychological stress, full bladder, eating a meal, exercise, smoking, taken coffee within 30 minutes before BP assessment were confounding factors based on less than 10% difference of ***r***, ICC, Kappa and BA results. The key results were as per Table 3 and 4 below.

Table 3: Pearson (r), Intra-class coefficient(ICC) and regression equations for BP methods

BP methods in comparison	Pearson coefficient(r) (*interpretation)	Intra-class Coefficient (†interpretation)	Regression equations for Relationship between BP methods (‡interpretation)
SYSTOLIC BP			
Standard/pragmatic	0.9(good association)	0.8(almost perfect)	SBPMc =-10.7+ 1.2SBPPr (gradient1.2,intercept10.7)
Standard/wrist	0.9(good association)	0.9(almost perfect)	SBPMc = 20 + 0.8SBPWr (gradient 0.8, intercept 20)
Pragmatic/wrist	0.9(good association)	0.9(almost perfect)	SBPPr= -2.5 + 1.0SBPWr (gradient 1.0, intercept -2.5)
DIASTOLIC BP			
Standard/pragmatic	0.9(good association)	0.9(almost perfect)	DBPPr = -0.7 + 1.0 DBPMc(gradient 1; intercept -0.7)
Standard/wrist	0.9(good association)	0.9(almost perfect)	DBPMc= 10.6+0.8DBPWr (gradient 0.8;intercept10.6)
Pragmatic/wrist	0.9(good association)	0.9(almost perfect)	DBPPr = 2.5 + 1.0DBPWr (gradient 1; intercept 2.5)

**Interpretation based on: -1.0 to - 0.7 strong negative association; -0.7 to -0.3 weak negative association; -0.3 to +0.3 little or no association; +0.3 to +0.7 weak positive association; +0.7 to +1.0 strong positive association.*
†Interpretation based on: ICC can be interpreted as follows: 0-0.2 indicates poor agreement: 0.3-0.4 indicates fair agreement; 0.5-0.6 indicates moderate agreement; 0.7-0.8 indicates strong agreement; and >0.8 indicates almost perfect agreement.
*‡Abbreviations: **SBPMc**-Sytolic BP Mercury; **SBPPr**-systolic BP Pragmatic; **SBPWr**-systolic BP wrist; **DBPPr**-diastolic BP wrist; **DBPMc**-diastolic BP Mercury; **DBPWr**-diastolic BP wrist.*

Table 4: Bland Altman analyses results and interpretation

	Bias (95% CI)	LIMITS OF AGREEMENT Lower (95% CI)	Upper (95% CI)	*DO THE METHODS CLINICALLY AGREE ?
SYSTOLIC BP				
Pragmatic/Ideal	-9.6 (-13.2 to -6.1)	-36.6 (-42.7 to -30.5)	17.4 (11.2 to 23.5)	No
Wrist/ideal	7.1 (4.1 to 10.0)	-15.4 (-20.5 to -10.3)	29.6 (24.5 to 34.7)	No
Pragmatic/wrist	-2.6(-5.8 to 0.7)	-26.9 (-32.4 to -21.4)	21.8 (16.3 to 27.3)	No
DIASTOLIC BP				
Pragmatic/Ideal	-3.0 (-5.6 to -0.4)	-22.6 (-27.0 to -18.1)	16.6 (12.1 to 21.0)	No
Wrist/Ideal	3.7 (1.6 to 5.7)	-11.7 (-15.1 to -8.2)	19.0 (15.5 to 22.5)	No
Pragmatic/Wrist	0.7 (-1.8 to 3.2)	-18.4 (-22.7 to -14.1)	19.8 (15.4 to 24.1)	No

***Interpretation based on**: comparison of limits of agreement to clinically acceptable range of BP, within 10 mmHg for diastolic BP and within 20 mmHg for systolic BP.*

***Comparison of pragmatic and standard BP**:* For systolic BP, the regression relationship was summarized as *SBPSMc = -10.7 + 1.2 SBPPg*. For agreement, the bias was 9.6 mmHg with limits of agreement of -17.4

mmHg to 36.6 mmHg. Using the bias alone, 9.6 mmHg, this would equate to excellent clinical inter-changeability based on clinically significant BP range of within 20 mmHg. However, the limits of agreement were too wide for the two methods to be regarded as clinically agreeing. Figure 1 illustrates the distribution on BA plot. For diastolic BP, the regression equation *DBPPr = -0.7 + 1.0 DBPMc* summarized the relationship of diastolic BP between pragmatic and standard mercury based BPs. The BA bias of 3.0 could have meant excellent agreement but the limits of agreement were again too wide (-16.6 mmHg to 22.6 mmHg) for agreement based on comparison to clinically interchangeable BP range of within 10 mmHg.

Comparison of wrist and mercury BP*:* For systolic BP, the corresponding regression equation was *SBPMc = -2.5 + 1.0SBPWr* . The BA analysis showed a bias of 7.1 mmHg and limits of agreement, -15.4 mmHg (lower) and 29.6 mmHg (upper), which were outside the clinical reference range for inter-changeability, within 20mmHg. For diastolic regression equation was linear, *DBPMc = 10.6 + 0.7DBPWr*, a sign of good positive association. The limits of agreement, -19.0 mmHg (lower) to 11.7 (upper), confirmed poor clinical agreement when compared to clinically acceptable range of agreement, within 10mmHg. Figure 2 below illustrates the regression line and BA plots.

Comparison of wrist and pragmatic BP*:* Finally pragmatic BP and wrist based standard BP were also compared for completeness. For systolic BP, ***r*** had a positive association. The BA plot, Fig. 3, show that the two methods could not be used interchangeably because the limits of agreement were wider than the within 20 mmHg clinical reference range. Similarly, for diastolic BP, the limits of agreement precluded exchangeable use as they were outside the within 10 mmHg clinical reference range.

Figure 1: BA plot for systolic and diastolic BP: standard mercury compared with pragmatic BP

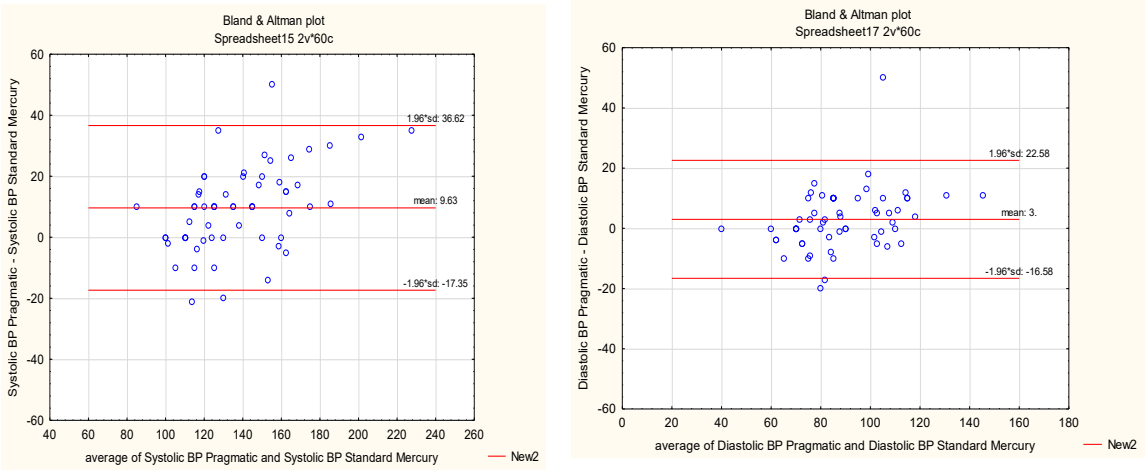


Figure 2: BA plots for systolic and diastolic BP: standard mercury compared with wrist BP

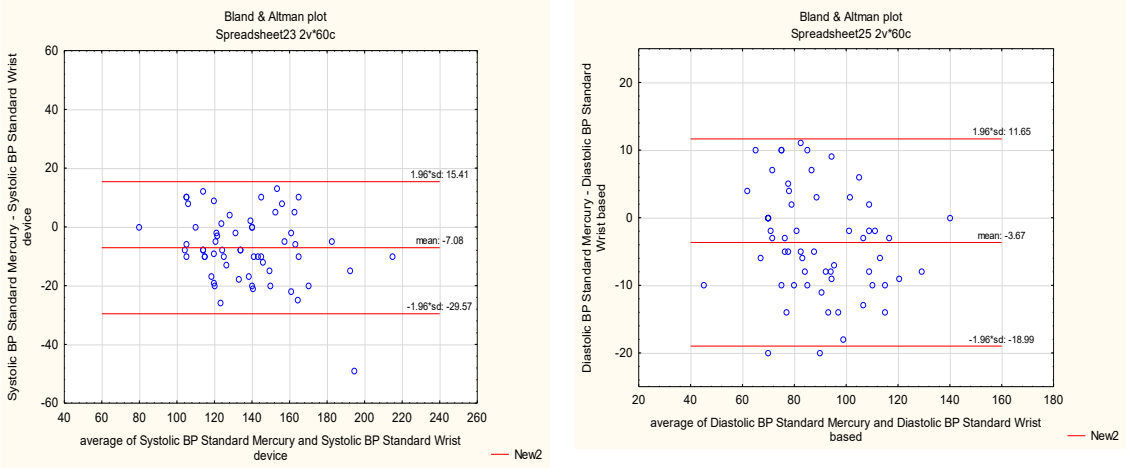
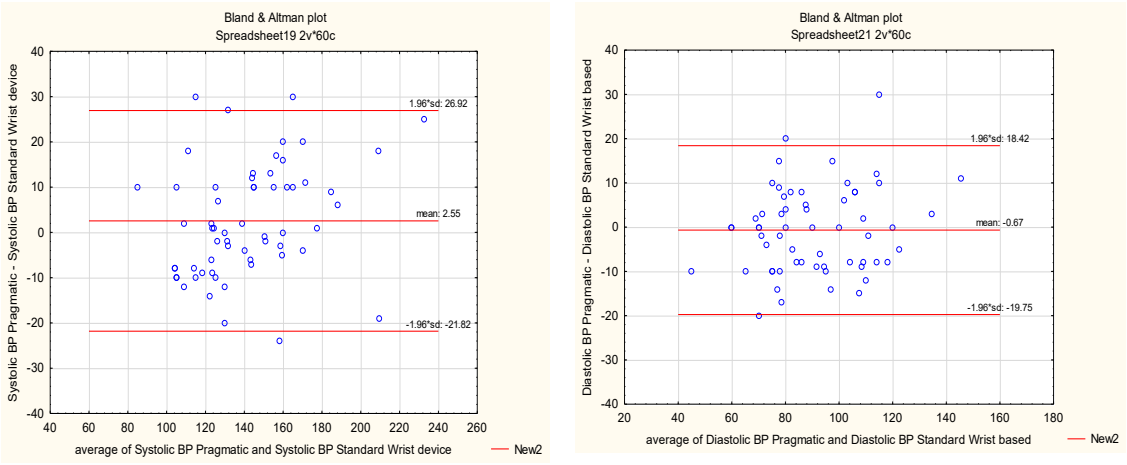


Figure 3: BA plots for systolic and diastolic BP: standard wrist compared with pragmatic BP



Comparison of treatment decisions: Treatment decisions (whether to start anti-hypertensive (1), alter anti-hypertensive treatment (2) or defer treatment (0)) were subsequently compared between decisions based on pragmatic BP and those based on standard mercury based BP. The Kappa score was 0.7 which equates to "good agreement" based on widely accepted Byrt's criteria (*see subscript under Table 6*). Overall (without stratifying), the treatment outcomes concurred in 83.8%, hence 16.7% were misclassified of treatment outcome when compared with the standard BP. For decision not to start treatment, 78% of instances concurred; for decision to start treatment, 90.9% agreed and for decision to adjust treatment, the agreement was 100%. In patients in whom treatment was deferred (basing on the standard mercury based BP), 19.5% (eight out of forty-one), were erroneously classified as requiring anti-hypertensive therapy using pragmatic

BP. Of those who needed to change treatment, the two BPs concurred (100%). Table 5 summarises the overall agreement level and Table 6 gave the stratified treatment outcomes.

Table 5: Comparison of treatment decisions between pragmatic based and standard based BP

Agreement	Expected Agreement	Kappa	95% CI Lower	95% CI Upper	p-value
83.8	44.2	0.7	0.5	0.9	0.0

Byrt criteria: Excellent agreement: 0.93-1.00; Very good agreement 0.81-0.92; good agreement: 0.61-0.80; fair agreement: 0.41-0.60; slight agreement: 0.21-0.40; poor agreement: 0.01-0.20; no agreement: <0.00

Table 6: Contingency table for per-stratum treatment outcomes comparing pragmatic to standard BP

		Treatment Plan based on Standard Mercury BP			
		0	**1**	**2**	**Total**
Treatment plan based on Pragmatic BP	**0**	32 (78.1%)	0 (0%)	0 (0%)	*32*
	1	8 (19.5%)	10 (90.9%)	0 (0%)	*18*
	2	1 (2.4%)	1 (9.1%)	8 (100.0%)	*10*
	Total	*41*	*11*	*8*	*60*

0 = no treatment; 1 = treatment; 2 = treatment changed

Summary of key findings from this study

1. The prevalence of hypertension was 25%.
2. There were inherent differences between standard mercury BP compared with standard wrist based BP; and between pragmatic BP compared with standard mercury based BP but not between pragmatic and wrist BP.
3. Pearson and intra-class correlation were similar for both systolic and diastolic BP and for all BP measurement pairs which were being compared, hence did not add much value in the comparisons.
4. BA analyses showed that pragmatic and standard BP measurement were different and could not be used inter-changeably in hypertension management. Standard mercury and wrist based methods were also not clinically inter-changeable.
5. Overall, treatment decisions between pragmatic BP based and standard BP based agreed in 83.3%. However, 16.7% of participants had their treatment outcomes misclassified.
6. Twenty percent of patients were started on anti-hypertensive therapy based on pragmatic BP when they actually did not need any treatment.
7. The decision to alter treatment was not affected by the differences between pragmatic and standard BP.
8. Ninety-one percent of those who truly needed treatment got treated when using pragmatic BP readings.

DISCUSSION

With hypertension defined as BP 140/90mmHg, one in five (20%) South Africans have hypertension, [4] a prevalence which is lower than 25% from this study. Since hypertension is more common in black ethnic race than other races, [4] this higher prevalence was likely due to the black-ethnic predominance of the study population.

In a study on BP measurement behaviour of clinicians done by Villegas et al, none of the physicians tested followed all the recommendations of the American Heart Association when measuring blood pressure, and a few recommendations were only followed by a minority of physicians. [17] In this study, we compared pragmatic and standard BP measurement. There was no agreement which could justify clinical inter-changeability between pragmatic and standard BP for both systolic and diastolic BP. Pragmatic systolic BP was at least 10.7 mmHg higher than standard mercury BP. For diastolic BP, pragmatic BP was at least 3 mmHg higher than the standard mercury readings. These results confirmed results from a study by Myers et al, which found that when the primary care physician records BP using a mercury or aneroid device, the resulting value frequently tends to be higher than what it would be if measurement guidelines were strictly adhered to. [18] Similarly Campbell et al concluded that pragmatic readings— those obtained with little attention to patient factors or recommended technique— cause errors in blood pressure assessment and are not highly correlated with target organ damage and as such, no evidence exists to support the use of pragmatic readings in assessing a patient's need for pharmacologic treatment. [19] Conversely, standardized readings— those that follow recommended protocols —correlate with hypertensive target organ damage and were used in the major randomized controlled trials that showed the benefits of pharmacotherapy. [19] The consequences are well documented in literature: consistent overestimation of diastolic BP by 5 mm Hg may more than double the number of patients with hypertension in a physician's practice. [20] People who are identified incorrectly as having hypertension may experience adverse effects of medication and have increased medical insurance and treatment costs. [21] Conversely, consistent underestimation of diastolic pressure by 5 mm Hg would reduce by 62% the number of patients perceived as hypertensive. [21] These errors could deprive patients of therapy proven to be beneficial and could lead to increases in serious medical and social complications. [21]

Comparison of wrist and mercury BP measurements was subsequently performed. Standard mercury diastolic and systolic BPs were consistently higher when using a wrist device. For systolic BP, the difference was as much as 20 mmHg and 10 mmHg for diastolic BP, a sharp contrast to previous studies which found similarities between mercury and wrist devices. [3,7,10,22] We suspected the difference was mostly due to the precise arm position and a known problematic phenomenon of wrist devices in which there is a systematic error introduced by the hydrostatic effect of differences in the position of the wrist relative to the heart. [22] This can be avoided if the wrist is always at heart level when the readings are taken, but there is no way of knowing retrospectively whether this was performed when a series of readings are reviewed. [22]

The variations discussed above had adverse effects on treatment decisions. Twenty-five percent of patients were started on anti-hypertensive therapy based on pragmatic BP when they actually did not need any treatment. This trend was similar to many studies which showed increased diagnosis of hypertension if blood pressure was not measured according to guidelines. [5,6,10,11,17,19,20] Overall, 16.7% of participants had their treatment outcomes misclassified. Of those who really needed treatment, there was a concordance of 91%. However, for those hypertensive patients who needed to have their treatment adjusted, pragmatic and standard BP had 100% concordance - the likely explanation being that when BP is grossly elevated, there is no difference between pragmatic and standard BP. One short fall of this use of misclassification as done here is that it does not differentiate between low magnitude inaccuracy; for example a BP of 89mmHg being misclassified as >90mmHg, which may be reasonably expected from any test, and high magnitude inaccuracy. A study which includes many BP values falling close to the defined cut off (of which hypertension has lots of cut off points for both diastolic and systolic) would show higher rates of misclassification between the methods being compared than a study where the majority of values lie away from the threshold. The other problem was diagnostic on the part of clinicians: the clinical decision to start treatment. Usually, a number of readings are required to start treatment unless there are risk factors, significant target organ damage or BP is significantly increased. The nurse practitioners might have over-diagnosed hypertension as they erroneously relied on one reading even when BP was mildly elevated.

The mercury sphygmomanometer is generally regarded as the gold standard against which all other devices for blood pressure measurement should be compared. [5] However recent studies have shown that ambulatory

BP measurements correlate better with the exact BP. Hodgkinson et al have recently concluded that ambulatory BP was more cost effective than clinic or home BP. [23] However guidelines for diagnosis and treatment of hypertension are still based on clinic BP measurements. [4,10] Statistical methods for comparison methods have been subject of discussion amongst clinicians. BA method is regarded as the gold standard. [24] Several papers have challenged the shortfalls of BA analysis. [25] On the other hand Bland and Altman, stated that the use of correlation coefficients is wrong for these types of studies. [24] In this study, intra-class correlation, Pearson's coefficient and linear regression both fell short of explicitly analyzing the research question.

Strength and weaknesses of the study

The main strength was that this study design was fast and inexpensive and was done in a resource limited setting like most third world institutions. It gave a useful initial overview of the problem including the community prevalence. The statistical methods used were appropriate for studies of this nature.

There was very limited potential to make causal inference of any differences, an obvious weakness of this study. Secondly, we could not claim success with minimising regression to mean with the serial BP measurements as the exact time to ensure regression to mean is rectified is unknown. In addition, it was impossible to totally eliminate observer bias despite blinding the nurses as there was always room for discussion when they meet outside the study centre hence the "pragmatic" BPs might not have been as pragmatic as we expected. Finally our gold standard was based on clinic mercury BP as opposed to the current recommendation, ambulatory BP.

What is already known on this topic?

There are differences between pragmatic and standard BP but wrist and mercury BP readings are usually comparable.

What this study adds.

This study further confirmed the existence of differences between pragmatic and standard BP measurements in resource limited setting. The difference leads to 16.7% disease status misclassification.

Wrist and mercury devices potentially lead to conflicting results contrary to earlier studies. Pearson and Intra-class correlation coefficients are weak statistical methods in studies of this nature.

Conclusion

There is a difference between pragmatic and standard BP measurements which affect decision to start treatment, decision todefer treatment but not treatment alteration decision for those already on treatment. There are also marked differences between wrist and standard mercury based BP devices which also affect treatment decision making.

Correlation of pragmatic BP with the new recommendation for use of ambulatory BP needs to be done as current guidelines are based on clinic BP-that would be more valuable. In future, when assessing agreement between clinical methods, BA method is more conclusive than correlation coefficients. Clinicians need to revert to basic good practice and measure BP more accurately to avoid unnecessary additional costs and morbidity associated with incorrect treatment due to disease misclassification. Wrist devices need to be used with caution.

Recommendations and dissemination of recommendation

Clinicians need to revert to basic good practice and measure BP more accurately to avoid unnecessary additional costs and morbidity associated with incorrect treatment due to disease misclassification. Contrary to existing research, wrist devices need to be used with caution.

These invaluable research findings will be disseminated to health professionals through a Powerpoint presentation and on-going BP measurement supervision both at RSSC and nationally. In addition, the findings will be published in an accredited peer-reviewed Family Medicine journal and conference presentation using posters and slides.

REFERENCES

1) Neaton JD, Wentworth D. Serum cholesterol, blood pressure, cigarette smoking, and death from coronary heart disease. Overall findings and differences by age for 316,099 white men. Multiple Risk Factor Intervention Trial Research Group. Arch Intern Med. 1992; 152: 56–64

2) Kim JW, Bosworth HB, Voils CI, Olsen M, Dudley T, Gribbin M, Adams M, Oddone EZ.How well do clinic-based blood pressure measurements agree with the mercury standard? J Gen Intern Med. 2005 Jul;20(7):647-9.

3) Angeli F, Sardone M, Angeli E, Repaci S, Gattobigio R, Verdecchia P.Validation of the A&D wrist-cuff UB-511 (UB-512) device for self-measurement of blood pressure. Blood Press Monit. 2006 Dec;11(6):349-54.

4) Joint National Hypertension Guideline Working Group 2006. South African Hypertension Guideline 2006. S Afr Med J 2006; 96: 337-362.

5) Jones DW, Appel LJ, Sheps SG, Roccella EJ, Lenfant C. Measuring blood pressure accurately: new and persistent challenges. Am Med Assoc JAMA. 2003;289:1027-1030.

6) Schwartz AR, Haas DC, Gerin W, Pickering TG. Accurate Measurement of Blood Pressure. JAMA. 2003; 289(21):2792.

7) Alexis O. Providing best practice in manual blood pressure measurement. Br J Nurs. 2009 Apr 9-22; 18(7):410-5. Review. PMID: 19373184 AD.

8) Wedgbury K, Valler-Jones T. Measuring blood pressure using an automated sphygmomanometer. British Journal of Nursing 17(11): 714 - 718 (12 Jun 2008).

9) Rose G. Standardisation of observers in blood pressure measurement. Lancet. 1965; 10: 673–674.

10) British Hypertension Society. Recommendation on blood pressure measurement. British Medical Journal Volume 293. 6 September 1986.

11) Roubsanthisuk W, Wongsurin U, Saravich S, Buranakitjaroen P. Blood pressure determination by traditionally trained personnel is less reliable and tends to underestimate the severity of moderate to severe hypertension. Blood Press Monit. 2007 Apr;12(2):61-8.PMID: 17353647.

12) McKay DW, Campbell NRC, Parab LS, Chockalingam A, Fodor JG. Clinical assessment of blood pressure. J Hum Hypertens 1990;4: 639-45.

13) Sala C, Santin E, Rescaldani M, Magrini F. How long shall the patient rest before clinic blood pressure measurement? American Journal of Hypertension, 2009 Volume 19, Issue 7, Pages 713-717.

14) Jhalani, Juhee, Tanya G, Lynn C, et al (2005). "Anxiety and outcome expectations predict the white-coat effect". Blood Pressure Monitoring 10 (6): 317–9. doi:10.1097/00126097-200512000-00006. PMID 16496447.

15) Altman DG, Bland JM. Frequently asked questions on the design and analysis of measurement studies: How to decide the sample size for a study of agreement between two methods of measurement. Website: www-users.york.ac.uk. Downloaded 30 November 2010.

16) Bland M. Regression towards mean or why was Terminator 3 such a disappointment. Department of Health. University of York. Website: www-users.york.ac.uk/~mb55/talks/regmean.htm.

17) Villegas I, Arias IC, Botero A, Escobar A. Evaluation of the technique used by health-care workers for taking blood pressure. Hypertension 1995;26:1204-6.

18) Martin G. Myers; Marshall Godwin; Martin Dawes; Alexander Kiss; Sheldon W. Tobe; Janusz Kaczorowski. Measurement of Blood Pressure in the Office Recognizing the Problem and Proposing the Solution; Hypertension. 2010 Feb; 55(2):195-200. Epub 2009 Dec 28.

19) Campbell RC, McKay WD. Accurate blood pressure measurement: Why does it matter? CMAJ. 1999 August 10; 161(3): 277–278.

20) Haynes RB, Sackett DL, Taylor DW, Gibson ES, Johnson AL. Increased work absenteeism from work after detection and labelling of hypertensive patients. N Engl J Med 1978;299:741-4.

21) Joffres MR, Hamet P, Rabkin SW, Gelskey D, Hogan K, Fodor G. Prevalence, control and awareness of high blood pressure among Canadian adults.CMAJ 1992;146(11):1997-2005.

22) Wonka F, Thummler M, Schoppe A. Clinical test of a blood pressure measurement device with a wrist cuff. Blood Press Monit. 1996; 1: 361–366.

23) Hodgkinson J, Mant J, Martin U, Hobbs FDR, Deeks JJ, Heneghan C, Roberts S, McManus RJ. Relative effectiveness of clinic and home blood pressure monitoring compared with ambulatory blood pressure monitoring in diagnosis of hypertension: systematic review. BMJ 2011; 342 doi: 10.1136/bmj.d3621 (Published 24 June 2011).

24) Bland M, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 1986; i: 307-310.

25) Pollock MA, Jefferson SG, Kane JW, MacKinnon G, Winnard CB. Method comparison – a different approach. Ann Clinical Biochemistry 1992;29:556 – 560.