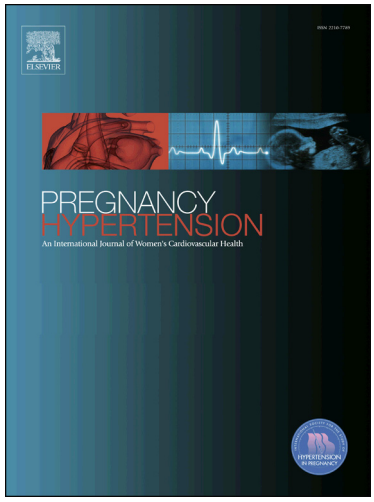


An inaccurate automated device negatively impacts the diagnosis and treatment of gestational hypertension

Diane Nzelu, Felicia Yeung, Delphine Couderq, Andrew Shennan, Nikos A. Kametas

PII: S2210-7789(17)30001-6
DOI: http://dx.doi.org/10.1016/j.preghy.2017.05.001
Reference: PREGHY 294

To appear in: *Pregnancy Hypertension: An International Journal of Women's Cardiovascular Health*

Received Date: 2 January 2017
Revised Date: 3 April 2017
Accepted Date: 1 May 2017

Please cite this article as: Nzelu, D., Yeung, F., Couderq, D., Shennan, A., Kametas, N.A., An inaccurate automated device negatively impacts the diagnosis and treatment of gestational hypertension, *Pregnancy Hypertension: An International Journal of Women's Cardiovascular Health* (2017), doi: http://dx.doi.org/10.1016/j.preghy.2017.05.001

An inaccurate automated device negatively impacts the diagnosis and treatment of gestational hypertension.

Dr Diane NZELU[1], MD
Miss Felicia YEUNG[1], BSc
Miss Delphine COUDERQ[1], MD
Professor Andrew SHENNAN[2], MD
Dr Nikos A KAMETAS[1,3], MD

[1] Antenatal Hypertension Clinic
Division of Women's Health
King's College Hospital, Denmark Hill
London SE5 9RS, UK

[2] Division of Women's Health
Department of Obstetrics and Gynaecology
St Thomas' Hospital,
London, SE1 7EH, UK

[3] Harris Birthright Research Centre for Fetal Medicine
Division of Women's Health
King's College Hospital, Denmark Hill
London SE5 9RS, UK

Corresponding Author

Nikos Kametas, MD, MRCOG
Harris Birthright Research Centre for Fetal Medicine
King's College Hospital, Denmark Hill
London, SE5 9RS
Email: nick.kametas@kcl.ac.uk
Tel: + 44 20 3299 2465
Fax: + 44 20 7738 3740

Word count (abstract): 247

Word count (main article): 3078

Abstract

Objectives

Automated blood pressure devices are frequently introduced in maternity care without prior validation for their accuracy in pregnancy. Our objectives were to, firstly, establish the accuracy in pregnancy of a locally used device (Welch Allyn 300) and, secondly, to audit its impact on the diagnosis and treatment of hypertension.

Study design

Validation study

The device was evaluated using the grading criteria of the European Society of Hypertension International Protocol (ESH-IP) (2010). Two observers took nine same-arm measurements alternating between the Welch Allyn and the mercury sphygmomanometer. Thirty-three women of any gestation were included.

Clinical audit

One observer took three same-arm measurements alternating between the Welch Allyn and the mercury sphygmomanometer. One hundred women of any gestation referred with suspected hypertension were included. The main outcome measures were the proportion diagnosed with hypertension or commenced on anti-hypertensive treatment on the presenting visit when using either the manual or the automated device.

Main Outcome Measures

Grading criteria of the ESH-IP (2010) and proportion of women diagnosed with hypertension or commenced on antihypertensive therapy at the presenting visit when using either manual sphygmomanometry or the Welch Allyn device.

Results

The Welch Allyn 300 series failed to meet the criteria of the ESH-IP (2010) for pregnancy. Compared to the mercury device, it under diagnosed hypertension by 48% and need for treatment by 80%.

Conclusions

The Welch Allyn 300 cannot be recommended for the measurement of blood pressure in pregnancy. Its use leads to the under-diagnosis and under-treatment of gestational hypertension.

Keywords: Blood pressure; hypertension; oscillometric; preeclampsia; pregnancy; validation; Welch-Allyn Vital Signs

Introduction

Hypertensive disorders affect approximately 10 percent of women during pregnancy and are the second direct cause of maternal death worldwide [1]. The accurate measurement of blood pressure is integral to the diagnosis and management of hypertension in pregnancy.

The mercury sphygmomanometer, the gold standard for the measurement of blood pressure, is becoming obsolete due to concerns regarding mercury toxicity [2, 3]. Automated devices are increasingly replacing mercury but their accuracy is disputed due to the empirical non-standardised algorithms used [4]. In attempt to address these concerns, validation protocols have been introduced and refined since the early 1990s [5-8]. The more recent version of the European Society of Hypertension International Protocol (ESH-IP) tightened the requirements for validation as a reflection of improvements in technology [7].

National guidelines do not state which method of blood pressure measurement should be used in pregnancy but only that it should be accurate and, if not mercury sphygmomanometry, validated [9]. To date, only 15 automated devices are validated for pregnancy of which five are accurate in preeclampsia [10-12]. Unfortunately, many hospitals use a centralised procurement system for blood pressure monitors and little attention is paid to their accuracy in different populations. At King's College Hospital, the Welch-Allyn Vital Signs 300 series has been centrally obtained and distributed to all maternity areas.

The two fundamental aspects of antenatal management of hypertension in pregnancy are, firstly, to be able to identify when a pregnant women becomes hypertensive and, secondly, to identify whether the blood pressure is above a pre-determined threshold at which treatment is required. The impact of using an automated device that is inaccurate for use in pregnancy or preeclampsia has not been previously reported.

In this study, we used the ESH-IP 2010 to evaluate the accuracy of the Welch-Allyn Vital Signs 300 series in pregnancy and preeclampsia. In addition, as this device was already in clinical use in our unit, we sought to determine its impact on the diagnosis and decision for treatment of hypertension amongst women referred with suspected hypertension to our Maternal Assessment Unit (MAU). Finally, we evaluated the association between maternal characteristics and device agreement, in an attempt to define future strategies to improve the device performance.

Methods

Validation study

Procedure

Women of any gestation with an arm circumference of 25.3 to 34.4cm were recruited from the maternity areas at Kings College Hospital (London, UK). Participants provided written consent at the time of recruitment. Women with unclear Korotkoff sounds (N=5), arrhythmia (N=1), or if unable to provide informed consent (N=1), were excluded from the study. Thirty-eight participants were excluded due to blood pressure range completion (N=28) or range adjustment (N=10).

The validation procedure recommended by the ESH-IP (2010) was followed [6]. Two observers were trained in blood pressure measurements using a CD-ROM programme produced by the British Hypertension Society. An independent supervisor assessed both observers individually prior to commencing the study to ensure correct blood pressure measurement technique.

The two observers took nine sequential same-arm measurements from each participant, alternating between the mercury sphygmomanometer and the Welch-Allyn Vital Signs 300 series device, using a double-headed stethoscope (Littman Class II Stethoscope). The participant was

seated, legs uncrossed and arm supported at heart level. An independent supervisor would request the mercury measurement to be repeated if the inter-observer discrepancy exceeded 4mmHg. A period of 30 seconds to 1 minute was allowed between readings to avoid venous congestion.

The first measurement by the two observers using the mercury sphygmomanometer was the entry blood pressure and was used to categorise the participant into a blood pressure range as specified by the ESH-IP. The second measurement using the automated device was used to orientate the device to the participant. The subsequent seven measurements were used in the final analysis.

Analysis

The mean differences between the device and the observer measurements were calculated to determine the accuracy and a visual representation of this was provided in the form of Bland – Altman plots. Based on these results, the device will 'pass' only if it satisfies the validation requirements.

Audit on the diagnosis and treatment of hypertension

Procedure

This was a prospective observational audit, which was carried out in parallel with the validation. We measured blood pressure sequentially using the Welch-Allyn Vital Signs 300 series device, which was routinely used in the MAU, and a mercury sphygmomanometer. We aimed to assess the differences in the blood pressure readings between the two devices and the impact that this would have on the diagnosis of hypertension in pregnancy within the index visit. In addition, we

aimed at assessing if maternal somatometric and demographic variables could explain potential differences between the two devices.

One hundred and fifteen women of any gestation who were referred with suspected hypertension were recruited from the MAU at Kings College Hospital. Verbal consent was obtained from each participant. Fifteen women were excluded due to unclear Korotkoff sounds (N=5), arrhythmias (N=2), incomplete outcome data (N=4), or due to being unable to provide informed consent (N=4).

The entry blood pressure was the blood pressure reading that prompted the referral. Each participant then had three sequential same-arm blood pressure readings with an interval of 30 seconds to one minute. A trained observer alternated between the mercury sphygmomanometer (first and third reading) and the Welch-Allyn Vital Signs 300 series test device (second reading). Three readings were taken to allow for two comparisons between the device and the manual sphygmomanometer and, thus, help limit bias.

Analysis

The outcome of each pregnancy was followed using our Maternity Database. Hypertension in pregnancy as defined by the American College of Obstetricians and Gynaecologists is blood pressure equal or greater than 140/90mmHg on two occasions, four hours apart. Patients were classed as having chronic hypertension if this predated pregnancy or was prior to 20 weeks gestation. Hypertension occurring beyond 20 weeks gestation was defined as pregnancy-induced hypertension (PIH) or preeclampsia if in the presence of significant proteinuria (i.e. $\geq$ 0.3g of protein in a 24-hour urine collection) [13].

Based on the definitions stated, we determined the number of patients who would have been diagnosed with new-onset hypertension had the Welch-Allyn Vital Signs 300 series or the

manual readings been used. To address the second issue of when to initiate anti-hypertensive treatment, again we looked at the number of patients who would have commenced medication based on either the Welch-Allyn Vital Signs 300 series or the mercury sphygmomanometer blood pressure readings. We used two thresholds for commencing treatment: 150/100mmHg, as recommended by the National Institute of Health and Clinical Excellence (NICE) [13], and 140/90mmHg, if tighter blood pressure control was to be achieved [9].

Statistics

As with the validation study, mean differences between the observer and the device were calculated and plotted on a Bland – Altman plot.

The normality of the distribution of the data was assessed by the Kolmogoroff-Smirnoff test. Univariate and multivariate linear and logistic regression were used to estimate and compare the magnitude of the effects of age, ethnicity, body mass index (BMI), gestational age, and systolic or diastolic blood pressure level (as measured by the mercury) on the differences between the two devices.

Results

Validation Study

Study participants

Seventy-eight women were recruited in order to complete the blood pressure categories as specified in the ESH-IP (2010). Thirty-three women were included in the final analysis. The characteristics of the study participants along with the number of observer measurements within each blood pressure category are presented in Table 1.

Amongst our cohort, 24 percent (N=8) had normal pregnancy outcomes, 48 percent (N=16) had preeclampsia, 18 percent (N=6) had PIH and 9 percent (N=3) were known to have chronic hypertension and had not developed preeclampsia during their pregnancy.

Validation criteria

The validation criteria of Phase 1 of the ESH-IP (2010) and the results of the validation analysis are presented in Table 2. The Welch-Allyn Vital Signs 300 series did not fulfil the protocol requirements and so has failed Phase 1 and 2 of the ESH-IP (2010) both for systolic and diastolic blood pressure. Bland-Altman plots of the data are given for systolic and diastolic blood pressure (Figure 1). These plots illustrate the relationship between the Welch-Allyn Vital Signs 300 series device and the mercury sphygmomanometer readings for the better observer.

Audit on the diagnosis and treatment of hypertension

Study participants

One hundred participants with new onset hypertension referred to the MAU were included in this audit. Group characteristics are outlined in Table 3. Their pregnancies were followed up prospectively until delivery. Fourteen percent (N=14) of the participants had chronic hypertension and 86 percent (N=86) had developed new-onset hypertension beyond 20 weeks gestation. Of the women with chronic hypertension, one developed preeclampsia, while from the new onset hypertension group, 48 women developed PIH and 38 women preeclampsia.

Clinical impact results

Four of the 14 women with chronic hypertension and 17 of the 86 women with new-onset hypertension were found to be normotensive on the day of referral by both the Welch-Allyn Vital Signs 300 series and the mercury sphygmomanometer. Of the new-onset group they all

subsequently developed preeclampsia (N=8) or PIH (N=9).

The mercury sphygmomanometer, using either the first or the second reading, classified 80 percent of new-onset patients (N= 69) on the day of referral as hypertensives (chi-square 37.8, p<0.0001). If only the first manual reading is used, then 77 percent (N= 66) of patients would have been classed as having new-onset hypertension (chi-square 32.1, p<0.0001) versus 72 percent (N= 62) if only the second manual blood pressure reading was used (chi-square 25.4, p<0.0001). On the contrary, the Welch-Allyn Vital Signs 300 series device classified as new-onset hypertensives only 32.6 percent (N=28) of referred patients (Figure 2).

Similar discrepancies between the performances of the two devices were seen regarding the need for antihypertensive treatment. If treatment was to be initiated above a threshold of greater than or equal to 150/100mmHg, 44 (N= 44) and 9 (N= 9) percent of all patients (chronic and new onset) were identified as requiring anti-hypertensives by the mercury or the Welch-Allyn Vital Signs 300 series device, respectively (chi-square 29.6, p<0.0001). If treatment was to be initiated above a threshold of greater than or equal to 140/90mmHg, then 78 (N= 78) and 34 (N=34) percent of patients (chronic and new onset) were identified as requiring anti-hypertensives by the mercury or the Welch-Allyn Vital Signs 300 series device, respectively (chi-square 37.5, p<0.0001) (Figure 2).

Amongst the hypertensive women (N=79), 13 (16.5%), 11 (14%) and 11 (14%) of them developed severe hypertension, thrombocytopenia and liver dysfunction, respectively.

Regression analysis

Device disagreement both in systolic and diastolic blood pressure was greater with increasing mercury blood pressure readings. This relationship was present when either the first or second manual readings were used. The device difference for the first manual systolic blood pressure

(Device Difference = 44.1-0.4*SBP1, p<0.0001, R^2=0.32, Supplement Figure 1) and second manual systolic blood pressure (Device Difference = 35.4-0.3*SBP2, p<0001, R^2=0.27, Figure 3) were higher with increasing blood pressure values. Device disagreement in diastolic blood pressure was higher with increasing levels of blood pressure and maternal age both for the first (Device Difference = 19.1-0.3*DBP1+0.2*age, p<0.0001, R^2=0.31, Supplement Figure 1) and second manual reading (Device Difference = 22.0-0.4*DBP2, p<0001, R^2=0.3, Figure 3). All other associations with BMI, ethnicity, heart rate and gestational age were not significant predictors of the device disagreement.

Discussion

Main Findings

This study has demonstrated that the Welch-Allyn Vital Signs 300 series failed to reach the required standards of the ESH-IP (2010) for the measurement of blood pressure during pregnancy and preeclampsia. The underestimation of the device became even more pronounced with increasing levels of blood pressure and maternal age. This discrepancy reached as high as 28 and 44mmHg, for systolic and diastolic blood pressure, respectively. Furthermore, the clinical significance of using an automated device underperforming in pregnancy was demonstrated by a prospective observational audit. In our population, when assessed by the mercury sphygmomanometer, 80 percent of women, who were eventually diagnosed with pregnancy-related hypertension, were found to be hypertensive. In contrast, only 32 percent of these women were identified when using the Welch-Allyn device, an underestimation of 48 percent.

NICE Guidelines recommends treatment when blood pressure is 150/100mmHg or above [9] in order to avoid placental hypoperfusion resulting in fetal growth restriction. However, a recent randomised controlled trial (RCT) demonstrated no significant difference in perinatal outcomes between those who obtained tight (i.e. target diastolic blood pressure of 85mmHg) versus less-

tight blood pressure control. In contrast, women with less-tight blood pressure control were almost twice as likely to develop severe hypertension [14]. As the appropriate cut-off for initiating antihypertensive treatment is not yet established, we assessed the impact of using mercury or an automated device for either of those thresholds.

About 80 and 56 percent fewer women would have been commenced on antihypertensive medication if their management had been based on the readings from the automated device for a threshold for treatment of 150/100 or 140/90mmHg, respectively. Based on our findings, it is possible that using an automated device that has not been validated for use in pregnancy may contribute to, first, failure to recognise women at risk of preeclampsia and, second, failure to then control their blood pressure.

In our study, realising how much the Welch-Allyn underestimated, we felt it was unethical but to base the management of the referred pregnancies on the readings of the mercury device. Therefore, we cannot present differences in pregnancy outcomes between the two devices. However, a Cochrane meta-analysis [15] and a recent RCT [14] have demonstrated a doubling in the risk of maternal severe hypertension, a risk factor for intracranial haemorrhage,[16] when mild hypertension is not treated to levels below 140/90. Furthermore, the same RCT has shown a more than doubling risk of thrombocytopenia and impaired liver function with less tight BP control.[14] In our data, assuming the total cohort had it's management based on the Welch-Allyn, it could have translated to a more than doubling in the rates of severe pre-eclampsia, thrombocytopenia and impaired liver function from the ones observed.

Strengths and Limitations

Simultaneous same-arm readings are the ideal for comparing two methods of measuring blood pressure. As this is not achievable when comparing mercury to automated devices, two alternatives exist; simultaneous opposite-arm and sequential same-arm measurements. In the

first, the inter-arm difference will vary from one measurement to the next. In the latter, three measurements are taken alternating between the mercury and the automated device. By calculating the difference between the second automated reading and the first and third manual readings and using the most favourable difference in the final analysis, parity with the simultaneous same-arm measurements is restored [17]. This explains the rationale behind the ESH-IP (2010) and emphasises the strength in our audit methodology

Furthermore, we assessed the association of maternal demographic variables and device discrepancy for both first and second mercury readings. This provided additional reassurance that the relationship demonstrated between the inaccuracy of the device, degree of hypertension and maternal demographics was consistent.

Interpretation

Our results agree with previous studies, which validated predecessors of this model and recommended caution with its use in hypertensive patients, whether they were pregnant [18] or not [19]. However, using the ESH-IP (2010), our results do not support the use of this model neither in normotensive nor in hypertensive pregnancies. This could be a reflection of the more stringent validation criteria of the ESH-IP (2010) applied in our study or a change in the device functional algorithms with more recent models [20].

The under-diagnosis of hypertension as a result of under-reading by an automated device is in agreement with data from non-pregnant populations. Turner et al showed that the diagnosis of hypertension is extremely sensitive to systematic errors in the measurement of blood pressure. They demonstrated that approximately 50 and 30 percent of patients with diastolic hypertension would be missed if diastolic blood pressure was underestimated by only 5 or 3 mmHg,

respectively. Underestimation of systolic blood pressure by 3 or 5 mmHg, would lead to the under-diagnosis of 15 and 25 percent of patients with systolic hypertension, respectively [20].

In our study, the underestimation in measuring both systolic and diastolic blood pressure was most discernible at higher readings. Similarly, Braam et al demonstrated that the accuracy of most currently available automated devices decreased at increasing blood pressure levels. This was attributed to a combination of the increasing variability seen in hypertensive patients and the fluctuation during sequential measurements [21]. One can argue that at very high readings the diagnosis of hypertension will not be affected. It is at mild levels of hypertension that there will be a more pronounced clinical impact from the lack of agreement and, as a result, a delay in instigating any treatment. Lim et al, compared the impact of using a validated automated device on the diagnosis of hypertension. In their study, 60 and 40 percent of misclassification between the two devices was for the diagnosis of pre-hypertension and Stage-1 hypertension, respectively, while there was agreement in the classification of Stage-2 hypertension [22].

The impact of the use of automated devices on the diagnosis of hypertension and disease outcome is well established in non-pregnant populations. However, the studies evaluating its impact in pregnant populations are sparse and all involved devices that were not validated for use in pregnancy or preeclampsia. The discrepancy between automated monitors and mercury sphygmomanometers is not similar across all devices, as each company employs different algorithms for the signal acquisition and analysis [4, 23]. It is, therefore, important that each device is specifically validated across all patient groups and the clinical impact of its use is shown in prospective studies. This is the reason why in our study, not only did we validate the Welch-Allyn device but we also showed the impact that its use had for the classification of hypertension.

The impact of demographic variables on the disagreement between mercury and automated devices has been previously demonstrated in non-pregnant populations [22, 24]. Larger studies are needed to assess whether adding these in the device algorithms would improve the performance of automated monitors.

Conclusion

The Welch-Allyn Vital Signs 300 underestimates systolic and diastolic blood pressure and, consequently, the diagnosis of hypertension and need for treatment at the presenting visit. Further studies are needed on the impact of the above on pregnancy outcome.

Details of Ethics Approval

Cambridge Research Ethics Committee, 12/EE/0191

Declaration of interest

The authors declare no conflict of interest

Acknowledgement

The authors would like to acknowledge the support provided by the staff and patients of MAU, King's College Hospital, London, UK.

References

[1] L. Say, D. Chou, A. Gemmill, Ö. Tunçalp, A.-B. Moller, J. Daniels, A.M. Gülmezoglu, M. Temmerman, L. Alkema, Global causes of maternal death: a WHO systematic analysis, The Lancet Global Health 2(6) (2014) e323-e333.

[2] UNEP, <UNEP_Global_Mercury_Assessment_2002.pdf>, (2002

).

[3] The European Parliament and the Council of the European Union, Directive 2002/95/EC of the European Parliament and of the Council on the restriction of the use of certain hazardous substances in electrical and electronic equipment, in: Offical Journal of the European Union (Ed.) 2003, pp. 19-23.

[4] J.N. Amoore, Oscillometric sphygmomanometers: a critical appraisal of current technology, Blood Pressure Monitoring 17(2) (2012) 80-88.

[5] E. O'Brien, J. Petrie, W. Littler, M. de Swiet, P. Padfield, D. Altman, M. Bland, A. Coats, N. Atkins, The British Hypertension Society protocol for the evaluation of blood pressure measuring devices, Journal of Hypertension 11(suppl 2) (1993) 543-562.

[6] E. O'Brien, T. Pickering, R. Asmar, M. Myers, G. Parati, J. Staessen, T. Mengden, Y. Imai, B. Waeber, P. Palatini, W. Gerin, Working Group on Blood Pressure Monitoring of the European Society of Hypertension International Protocol for validation of blood pressure measuring devices in adults, Blood Press Monit 7(1) (2002) 3-17.

[7] E. O'Brien, N. Atkins, G. Stergiou, N. Karpettas, G. Parati, R. Asmar, Y. Imai, J. Wang, T. Mengden, A. Shennan, European Society of Hypertension International Protocol revision 2010 for the validation of blood pressure measuring devices in adults, Blood Press Monit 15(1) (2010) 23-38.

[8] American National Standards Institute, American national standard for electronic or automated sphygmomanometers, Association for the Advancement of Medical Instrumentation, (1987).

[9] National Institute for Health and Clinical Excellence, Hypertension in pregnancy: the management of hypertensive disorders during pregnancy, NICE guideline (CG107). 2010.

[10] A. de Greeff, Z. Beg, Z. Gangji, E. Dorney, A.H. Shennan, Accuracy of inflationary versus deflationary oscillometry in pregnancy and preeclampsia: OMRON-MIT versus OMRON-M7, Blood pressure monitoring 14(1) (2009) 37-40.

[11] A. Reinders, A. Cuckson, J. Lee, A. Shennan, An accurate automated blood pressure device for use in pregnancy and pre‐eclampsia: the Microlife 3BTO‐A, BJOG: An International Journal of Obstetrics & Gynaecology 112(7) (2005) 915-920.

[12] Y. Chung, M. Brochut, A. de Greeff, A. Shennan, Clinical accuracy of inflationary oscillometry in pregnancy and pre-eclampsia: Omron-MIT Elite, Pregnancy Hypertension: An International Journal of Women's Cardiovascular Health 2(4) (2012) 411-415.

[13] Task Force on Hypertension in Pregnancy, Hypertension in Pregnancy American Journal of Obstetrics and Gynaecology Practice Guideline WQ 244 (2013).

[14] L.A. Magee, P. von Dadelszen, E. Rey, S. Ross, E. Asztalos, K.E. Murphy, J. Menzies, J. Sanchez, J. Singer, A. Gafni, A. Gruslin, M. Helewa, E. Hutton, S.K. Lee, T. Lee, A.G. Logan, W. Ganzevoort, R. Welch, J.G. Thornton, J.M. Moutquin, Less-tight versus tight control of hypertension in pregnancy, N Engl J Med 372(5) (2015) 407-17.

[15] E. Abalos, L. Duley, D.W. Steyn, D.J. Henderson‐Smart, Antihypertensive drug therapy for mild to moderate hypertension during pregnancy, The Cochrane Library (2007).

[16] M. Knight, D. Tuffnell, S. Kenyon, J. Shakespeare, R. Gray, J.E.o.b.M.-U. Kurinczuk, Saving Lives, Improving Mothers' Care - Surveillance of maternal deaths in the UK 2011-13 and lessons learned to inform maternity care from the UK and Ireland Confidential Equiries into Maternal Deaths and Morbidity 2009-13., National Perinatal Epidemiology Unit, University of Oxford,, Oxford, 2015.

[17] N. Atkins, F. Mee, K. O'Malley, E. O'Brien, The relative accuracy of simulataneous same arm, simultaneous opposite arm and sequential same arm measurements in the validation of automated blood pressure measuring devices, Journal of Human Hypertension 4 (1990) 647-649.

[18] A. Reinders, A. Cuckson, C. Jones, R. Poet, G. O'Sullivan, A. Shennan, Validation of the Welch Allyn 'Vital Signs' blood pressure measurement device in pregnancy and pre-eclampsia, British Journal of Obstetrics and Gynaecology 110 (2003) 134-138.

[19] C. Jones, K. Taylor, L. Poston, A. Shennan, Validation of the Welch Allyn 'Vital Signs' oscillometric blood pressure monitor, Journal of Human Hypertension 15(3) (2001) 191-195.

[20] M.J. Turner, A.B. Baker, P.C. Kam, Effects of systematic errors in blood pressure measurements on the diagnosis of hypertension, Blood Press Monit 9(5) (2004) 249-53.

[21] R.L. Braam, T. Thien, Is the accuracy of blood pressure measuring devices underestimated at increasing blood pressure levels?, Blood Press Monit 10 (2005) 283.

[22] Y.H. Lim, S.Y. Choi, K.W. Oh, Y. Kim, E.S. Cho, B.Y. Choi, Y.M. Kim, J. Shin, Comparison between an automated device and a manual mercury sphygmomanometer in an epidemiological survey of hypertension prevalence, Am J Hypertens 27(4) 537-45.

[23] W. Gerin, A. Schwartz, J. Schwartz, T. Pickering, K. Davidson, J. Bress, E. O'Brien, N. Atkins, Limitations of current validation protocols for home blood pressure monitors for individual patients, Blood Pressure Monitoring 7 (2002) 313-318.

[24] Y. Ostchega, R.J. Prineas, T. Nwankwo, G. Zipf, Assessing blood pressure accuracy of an aneroid sphygmomanometer in a national survey environment, Am J Hypertens 24(3) 322-7.

Table 1: Validation study (ESH-IP 2010) - Subject details.

Subject details (N=33)			
Age (years)			
Range (Low:High)		22:41	
Mean (SD)		31.4 (5.1)	
Arm circumference (cm)			
Range (Low:High)		24:32.5	
Mean (SD)		28.9 (2.6)	
Recruitment Systolic BP (mmHg)			
Range (Low:High)		85:177	
Mean (SD)		136.9 (24.9)	
Recruitment Diastolic BP (mmHg)			
Range (Low:High)		45:110	
Mean (SD)		85.8 (18.1)	
Observer measurements in each recruitment range			
Systolic BP		Diastolic BP	
Overall range (low:high)	82:171	Overall range (low:high)	48:121
Low (<130)	55	Low (<130)	27
Medium (130-160)	38	Medium (130-160)	48
High (>160)	6	High (>160)	24
Maximum difference	49	Maximum difference	24

Table 2: Validation study – Grading criteria and results for the ESH-IP (2010)

Part 1 **Pass requirements**	<5mmHg	<10mmHg	<15mmHg	Grade 1
Two of	73	87	96	
All of	65	81	93	
Achieved				
SBP	41	66	84	Fail
DBP	18	59	81	Fail
Part 2 **Pass requirements**	2/3 ≤ 5mmHg	0/3 ≤ 5mmHg		Grade 2
All of	>24	<3		
Achieved				
SBP	13	14		Fail
DBP	4	22		Fail

Table 3: Clinical impact audit – Subject details

Subject details (N=100)	
Blood pressure status	
- Chronic hypertension, n	14
- Pre-eclampsia, n	38
- Pregnancy-induced hypertension, n	48
Age (years)	
Low:High	18:57
Mean (SD)	32.95(±6.4)
BMI (kg/m^2)	
Low:High	19:52
Mean (SD)	27.95 (±6.3)
Ethnicity	
Black/African, n	34
Black/Carribbean, n	17
White/Caucasian, n	45
South East Asian, n	4
Gestational age (weeks)	
Low:High	11:41.4
Mean (SD)	35.4 (±8.23)

Figure Legends

- **Figure 1A:**
  - **Title:** Difference of systolic blood pressure - observer
  - **Description:** Bland Altman plot showing the difference in systolic blood pressure between the Welch Allyn Vital Signs 300 device readings and the mean of two observer readings in all 33 participants.
- **Figure 1B:**
  - **Title**: Difference of diastolic blood pressure – observer
  - **Description:** Bland Altman plot showing the difference in diastolic blood pressure between the Welch Allyn Vital Signs 300 device readings and the mean of two observer readings in all 33 participants.
- **Figure 2A:**
  - **Title:** Diagnosis of hypertension.
  - **Description:** Proportion of patients classified as hypertensive on the index visit by the first, second or highest of the two mercury measurements (grey boxes) and by the Welch-Allyn Vital Signs 300 (black box).
  - **Abbreviations:** BP – blood pressure
- **Figure 2B:**
  - **Title:** Treatment of hypertension.
  - **Description:** Proportion of patients identified as needing treatment on the index visit by the first, second or highest of the two mercury sphygmomanometer measurements (grey boxes) and by the Welch-Allyn Vital Signs 300 (black box). On the left, if a cut off of 150/100 mmHg and on the right if a cut off of 140/90 mmHg was used to determine the need for treatment.
  - **Abbreviations:** BP – blood pressure

- **Figure 3A**
    - **Title:** Systolic blood pressure
    - **Description:** Scatter plot describing the relationship between the difference of the device and the second mercury reading (Y-axis) and the second mercury reading (x-axis). The regression line (95% CI) illustrating the above relationship is also shown.
    - **Abbreviations:** CI – confidence interval
- **Figure 3B:**
    - **Title:** Diastolic blood pressure
    - **Description**: (*Left)* Scatter plot describing the relationship between the difference of the device and the second mercury measurement (Y-axis) and the mercury second measurement (x-axis). The regression line (95% CI) illustrating the above relationship is also shown. *(Right)* Scatter plot describing the relationship between the difference of the device and the second mercury measurement (Y-axis) and the maternal age (x-axis). The regression line (95% CI) illustrating the above relationship is also shown.
    - **Abbreviations:** CI – confidence interval
- **Supplementary Figure 1A**
    - **Title:** Systolic blood pressure
    - **Description:** Scatter plot describing the relationship between the difference of the device and the first mercury reading (Y-axis) and the first mercury reading (x-axis). The regression line (95% CI) illustrating the above relationship is also shown.
    - **Abbreviations:** CI – confidence interval
- **Supplementary Figure 1B**
    - **Title:** Diastolic blood pressure
    - **Description**: (*Left)* Scatter plot describing the relationship between the difference of the device and the first mercury measurement (Y-axis) and the first second measurement (x-axis). The regression line (95% CI) illustrating the above

relationship is also shown. *(Right)* Scatter plot describing the relationship between the difference of the device and the first mercury measurement (Y-axis) and the maternal age (x-axis). The regression line (95% CI) illustrating the above relationship is also shown.

- **Abbreviations:** CI – confidence interval

Figure 1

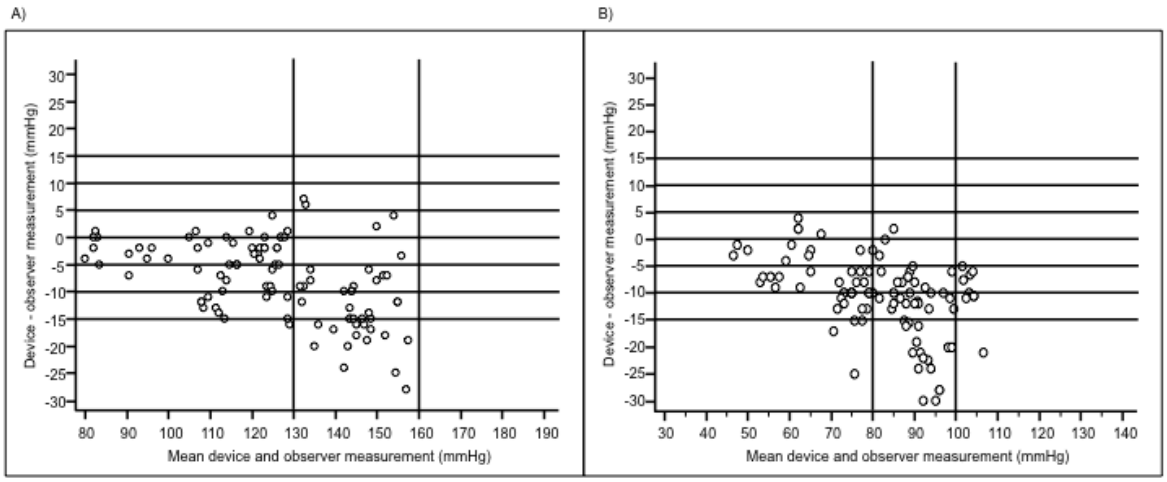


Figure 2

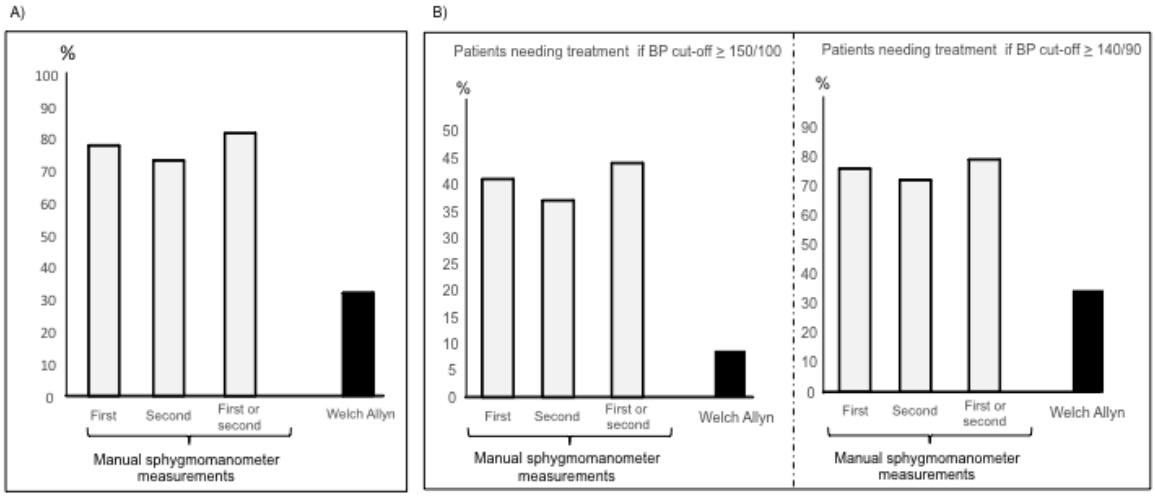
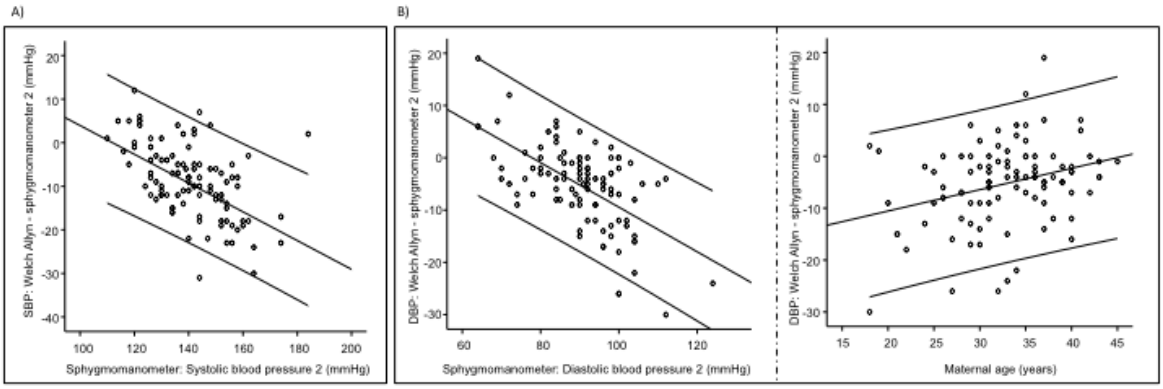


Figure 3



Highlights

An inaccurate automated device negatively impacts the diagnosis and treatment of gestational hypertension.

- Welch-Allyn Vital Signs 300 series underestimates blood pressure in pregnancy.
- The underestimation is more pronounced at higher blood pressures and maternal age.
- The Welch-Allyn device leads to a 48% under-diagnosis of gestational hypertension.
- Its underestimation results in under-treatment of gestational hypertension by 80%.