Pitfalls in blood pressure measurement in daily practice

ST Houweling\textsuperscript{a,b}, N Kleefstra\textsuperscript{a}, HL Lutgers\textsuperscript{a,c}, KH Groenier\textsuperscript{b}, B Meyboom-de Jong\textsuperscript{b} and HJG Bilo\textsuperscript{a}

Houweling ST, Kleefstra N, Lutgers HL, Groenier KH, Jong BM and Bilo HJG. Pitfalls in blood pressure measurement in daily practice. \textit{Family Practice} 2006; 23: 20–27.

Background. Accurate blood pressure (BP) readings and correctly interpreting the obtained values are of great importance. However, there is considerable variation in the different BP measuring methods suggested in guidelines and used in hypertension trials.

Objective. To compare the different methods used to measure BP; measuring once, the method used for a large study such as the UKPDS, and the methods recommended by various BP guidelines.

Methods. In 223 patients with type 2 diabetes from five family practices BP was measured according to a protocol to obtain the following data: A = first reading, B = mean of two initial readings, C = at least four readings and the mean of the last three readings with less than 15% coefficient of variation difference, D = mean of the first two consecutive readings with a maximum of 5 mm Hg difference. Mean outcomes measure is the mean difference between different BP measuring methods in mm Hg.

Results. Significant differences in systolic/diastolic BP were found between A and B [mean difference (MD) systolic BP 1.6 mm Hg, \( P < 0.001 \)], B and C (MD 5.7/2.8 mm Hg, \( P < 0.001 \)), B and D (MD 6.2/2.8 mm Hg, \( P < 0.001 \)), A and C (MD 7.3/3.3 mm Hg), and A and D (MD 7.9/3.0 mm Hg, \( P < 0.001 \)).

Conclusion. Different methods to assess BP during one visit in the same patient lead to significantly different BP readings and can lead to overestimation of the mean BP. These differences are clinically relevant and show a gap between different methods in trials, guidelines and daily practice.

Keywords. Blood pressure determination, clinical trials, human, hypertension, practice guidelines.

Introduction

In hypertension, accurate blood pressure (BP) readings and the correct interpretation of the obtained values are of great importance to epidemiology as well as to diagnosis, treatment and research.\textsuperscript{1,2} Although a large number of guidelines and recommendations describing how blood pressure BP measured are available, research shows that health care providers frequently do not comply with these guidelines.\textsuperscript{3–5} This leads to possible mistakes in the diagnosis and treatment of hypertension. Additionally, the guidelines are not always consistent with each other. For example, the European Society of Hypertension recommends calculating the mean of at least two BP readings on each visit.\textsuperscript{5} This method corresponds to the measuring methods described in the seventh report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of high blood pressure (the JNC7 report), the British Hypertension Society, the Dutch College of General Practitioners and Perloff et al.\textsuperscript{6–9} However, the Dutch Institute of Healthcare (CBO) recommends taking as many readings as are necessary to obtain two readings that are no >5 mm Hg apart (systolic or diastolic).\textsuperscript{10} The mean of these two readings is then considered to be an accurate representation of the patient’s BP. Furthermore (and potentially more disturbing when it comes to interpreting data) studies on antihypertensive medication or cardiovascular risk prediction use different methods to measure BP, and
some studies do not report which method was used. Table 1 shows that various protocols are used in these trials to measure BP and that these protocols do not always agree with the guidelines. The number of measurements varies from one to six, the resting period from 0 to 5 minutes, and the time interval between two measurements from 0 to 30 seconds. The BP cut off points in the guidelines, which are to be striven for in patients with hypertension, are usually based on these large clinical or epidemiological trials. For example, based on the United Kingdom Prospective Diabetes Study (UKPDS), until 2003, Dutch general practitioners aimed for a BP below 150/85 mm Hg in patients with diabetes mellitus type 2 (according to more recent guidelines, the target BP in these patients is 140/90 mm Hg). However, as we mentioned above, the guidelines used by the Dutch general practitioners describe a different method for measuring BP (mean of the first two readings) than was used in the UKPDS (4 readings, mean of the last three measurements).

The aim of our study is to compare the differences in BP readings resulting from applying the different methods in a single patient, during a single visit: the method used most widely in daily practice (measuring once), the methods recommended in various guidelines, and the method used in the UKPDS.

### Methods

#### Setting

In January 2003, all patients with diabetes mellitus type 2 registered in the group practice of five general practitioners were invited by letter to come in for their annual diabetes check-up. The annual check-up includes BP measurement. Two trained physicians were randomly assigned patients, and they followed a standard protocol when taking the BP readings. All the patients gave written informed consent.

#### Equipment

A calibrated OMRON M5-I (HEM-757) automatic blood-pressure device was used. This device was validated according to different international protocols.

#### BP measurements

Before taking the blood pressure, the circumference of the upper arm was measured. When the circumference

---

### Table 1  Blood pressure measuring protocols as described in some large trials on antihypertensive therapy

<table>
<thead>
<tr>
<th>Trial</th>
<th>References</th>
<th>Device</th>
<th>Mercury/aneroid/automatic</th>
<th>Method (auscultatory/oscillometric)</th>
<th>Measurement protocol</th>
<th>Blood pressure value used for the study</th>
</tr>
</thead>
<tbody>
<tr>
<td>UKPDS</td>
<td>22,23</td>
<td>Copal UA.251 or</td>
<td>Automatic</td>
<td>Auscultatory</td>
<td>Sitting, at least 5 minutes rest, minimal 4 readings</td>
<td>Mean of 2nd, 3rd, and 4th reading with a coefficient of variation below 15%</td>
</tr>
<tr>
<td>ALLHAT</td>
<td>24,25</td>
<td>Takeda UA.751</td>
<td>Automatic</td>
<td>Auscultatory</td>
<td>2 readings (30 seconds interval)</td>
<td>Mean of both readings</td>
</tr>
<tr>
<td>LIFE</td>
<td>26–28</td>
<td>Hawksley RZS</td>
<td>Mercury</td>
<td>Auscultatory</td>
<td>Sitting, after 5 minutes rest</td>
<td>Not described</td>
</tr>
<tr>
<td>HOT</td>
<td>29</td>
<td>Visomat OZ D2</td>
<td>Automatic</td>
<td>Oscillometric</td>
<td>Sitting, after 5 minutes rest, 3 readings</td>
<td>Mean of all three readings</td>
</tr>
<tr>
<td>RENAAL</td>
<td>30, 31</td>
<td>Not described</td>
<td>Not described</td>
<td>Not described</td>
<td>Sitting, after 5 minutes rest, 3 readings</td>
<td>Mean of all three readings</td>
</tr>
<tr>
<td>ANBP2</td>
<td>32, 33</td>
<td>Not described</td>
<td>Mercury</td>
<td>Auscultatory</td>
<td>Sitting, 2 readings</td>
<td>Not described</td>
</tr>
<tr>
<td>HOPE</td>
<td>34, 35</td>
<td>Not described</td>
<td>Not described</td>
<td>Not described</td>
<td>Not described</td>
<td>Not described</td>
</tr>
<tr>
<td>IRMA2</td>
<td>36</td>
<td>Not described</td>
<td>Sphygmomanometer</td>
<td>Not described</td>
<td>Sitting, after 10 minutes rest, 2 readings (2 minutes interval)</td>
<td>Mean of both readings</td>
</tr>
<tr>
<td>Syst-Eur</td>
<td>37–39</td>
<td>Not described</td>
<td>Not described</td>
<td>Not described</td>
<td>Sitting, at randomisation: 6 readings; follow-up measurements: not described</td>
<td>Mean of six readings (at randomisation)</td>
</tr>
<tr>
<td>SHEP</td>
<td>40</td>
<td>Hawksley RZS</td>
<td>Mercury</td>
<td>Auscultatory</td>
<td>sitting, 2 readings</td>
<td>Mean of both readings</td>
</tr>
<tr>
<td>PROGRESS</td>
<td>41</td>
<td>Not described</td>
<td>Mercury</td>
<td>Auscultatory</td>
<td>2 readings</td>
<td>Not described</td>
</tr>
</tbody>
</table>
was 22–32 cm, the standard cuff (12 × 21.5 cm) was used. For circumferences between 32 and 42 cm, the large cuff (15 × 29.5 cm) was used. The measurement was done with the patient in a sitting position, after he or she had been sitting for a minimum of 5 minutes, in such a way that the cuff was at heart level and the volar side of the lower arm rested on the desk. The cuff was applied to the bare arm 1–2 cm above the elbow fold. Any tight clothing was removed from the upper arm. The patient was asked to sit still, not to move the arm, and not to speak during measurement. The time interval between successive measurements was at least 15 seconds. Initially, BP was measured twice and not to speak during measurement. The time of the lower arm rested on the desk. The cuff was applied to the bare arm 1–2 cm above the elbow fold. Any tight clothing was removed from the upper arm. The patient was asked to sit still, not to move the arm, and not to speak during measurement. The time interval between successive measurements was at least 15 seconds.10 Initially, BP was measured twice in each arm, with the choice of arm at the discretion of the patient. The first reading was recorded as such. The mean of the two readings for the left arm was compared with the mean for the right arm. When there was a difference of >10 mm Hg between the systolic and/or diastolic BP readings, the measurements were continued on the arm with the higher BP, which is according to the Dutch guidelines.8,10 When the difference was less, an arbitrary arm was taken for the next measurements. This arbitrary arm or the arm with the higher BP was used to obtain the mean of two BP readings. Again, two more readings were done, bringing the total number of readings for one individual in one arm to four. When a coefficient of variation above 15% (SD*100%/mean) in the last three consecutive measurements was found, additional readings were taken until the last three were below 15% (to comply with the UKPDS protocol). Furthermore, readings were taken until two readings ≤5 mm Hg apart were obtained for systolic and diastolic BP.

In this way, everybody complied with the four following measurement methods: first reading (Method A), mean of the first two readings; (Method B), at least four readings and the mean of the last three readings with <15% coefficient of variation difference; (Method C), the mean of the first two consecutive readings with a maximum of 5 mm Hg difference (Method D).

Data entry and analysis
We used SPSS to set up the database and analyse the data. All data showed a normal distribution. The mean BP readings obtained using the different methods were compared using a general linear model (GLM repeated measures) and we adjusted the outcome analyses using the Bonferroni correction. The differences between the two methods against their means were analysed in a Bland and Altman plot.17

Results
The results are presented in Table 2. Of the 287 patients invited to participate, 223 patients were included in the study population (78% response). All patients who did not participate in the study had either limited mobility, making a visit to the practice impractical, or were treated for their diabetes by an internist at the hospital. The average age of the participants was 68.5 years (range: 36–91 years). In 23% (n = 50) of the patients we had to use a large occluding cuff to measure the BP (mean arm circumference of the whole group being 29.8 cm). When applying the rule that BP should be measured until two readings are obtained that are ≤5 mm Hg apart, in 43% (n = 92) two readings sufficed. On average 3.5 readings were needed to reach this goal.

Table 3 presents a comparison of the different measuring methods. Nearly all the methods differed with regard to systolic BP (P < 0.001), except C and D. However, 34% (n = 72) of the patients showed a difference of >5 mm Hg between methods C and D. For diastolic BP similar inter-group differences were found (P < 0.001), except between A and B, and between C and D. Methods C and D resulted in the lowest BP readings.

Seventy-two percent (n = 161) of patients had a BP above 150/85 mm Hg, measured according to the UKPDS method (Method C). Table 4 shows that the other measuring methods were sensitive (all above 85%) in predicting this target UKPDS value, but only Method D was specific (86%). For any particular positive test result, the probability that it is a false positive (measuring a BP above 150/85 mm Hg while with the UKPDS method the BP was lower than 150/85 mm Hg) was 34, 29 and 14.5% according to Method A, Method B, or Method D, respectively.

The Bland and Altman plot analyses between the UKPDS method and the other methods are illustrated in Figure 1. These analyses demonstrate large differences between the different measurement methods. The regression lines drawn in the plots indicate that the differences increase as BP increases (except when comparing the BP between Method D and C).

Discussion
The method used to assess BP determines the level of the BP found. For example, if the method of the European Society of Hypertension is used, it results in systolic and diastolic BP readings that are, respectively, 5.7 and 2.8 mm Hg higher than when the UKPDS protocol is used. This effect is higher than the differences that are considered clinically relevant in large trials (a decrease in systolic BP of 5 mm Hg is considered significant).

Is this relevant for daily practice? The UKPDS-method is considered the gold standard in diabetes care. This implies that there is an overestimation (and possibly an over-treatment) of the BP in a significant number of patients (29%) if the method outlined by the
European Society of Hypertension is followed or if health care professionals are taking only a single reading (34%). This finding may have implications for the currently held opinion that BP control in type 2 diabetic patients is generally quite poor.18–21 If different measuring protocols are being used, then an artefact due to inadequate measurement may partially explain these perceptions.

We found the large differences in BP when we compared Method A with Method C. A difference between two methods is more or less irrelevant if the systolic BP is high (200 mm Hg or more). In contrast, a difference of 10 mm Hg has direct impact on treatment decisions if the systolic BP is close to the target value. We found that 99% of patients with a normal BP according to the UKPDS-method (<150/85) had a BP up to 167/92 mm Hg when it was measured only once. Health care professionals who are measuring once and who find a BP above 167/92 mm Hg are correctly qualifying the BP as too high in 99% of the cases in comparison with the UKPDS method, but the problem still stands: which target BP is appropriate?

We used, according to the Dutch guidelines, the highest blood pressure in cases of inter-arm blood pressure differences >10 mm Hg.8,10 It would be expected that if we used a random BP, the differences between methods would be somewhat less, because of lower average BP readings.

Another problem affecting daily practice is that the BP measurement recommendations, such as posture of subject, arm support, arm position and cuff size,3 are rarely followed, and the equipment used is sometimes inaccurate.5 Besides the differences in measuring methods between the various guidelines and the hypertension trials, the description of the method used in large hypertension trials is rather poor (Table 1).

Only four of the 11 studies in Table 1 supply a complete description of the method used, including the BP measuring device, the number of readings, the time between measurements, the period the patient waited before the first measurement and the way in which the final BP was calculated. In four studies the descriptions are incomplete, and three studies mentioned hardly any details. Even when there is an exact description, one often has to refer to the very first publication of the study in which the study design is described. These

<table>
<thead>
<tr>
<th>Patient characteristics</th>
<th>N</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td>48% (n = 107) male</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td>68.5 years (SD 11.3)</td>
</tr>
<tr>
<td>Upper arm circumference</td>
<td>204</td>
<td>29.8 cm (SD 3.3)</td>
</tr>
<tr>
<td>Number of patients with upper arm circumference &gt;32 cm</td>
<td>219</td>
<td>23% (n = 50)</td>
</tr>
<tr>
<td>Number of readings required to get 2 consecutive readings (systolic/diastolic) &lt;5 mm Hg apart</td>
<td>212</td>
<td>3.5 measurements (min 2, max 13) (SD 1.9)</td>
</tr>
<tr>
<td>Patients in which an x number of readings was sufficient to get two consecutive readings &lt;5 mm Hg apart</td>
<td>212</td>
<td>2 readings = 43.4% (n = 92)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 readings = 12.3% (n = 26)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4 readings = 21.7% (n = 46)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5 readings = 9.4% (n = 20)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6 readings = 6.1% (n = 13)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7 readings = 2.8% (n = 6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;7 readings = 4.3% (n = 9)</td>
</tr>
<tr>
<td>Mean blood pressure (systolic/diastolic) Method A</td>
<td>215</td>
<td>160.6/88.3 mm Hg (SD 23.4/11.9)</td>
</tr>
<tr>
<td>Mean blood pressure (systolic/diastolic) Method B</td>
<td>219</td>
<td>159.2/87.9 mm Hg (SD 22.5/11.1)</td>
</tr>
<tr>
<td>Mean blood pressure (systolic/diastolic) Method C</td>
<td>218</td>
<td>153.7/85.2 mm Hg (SD 21.4/10.0)</td>
</tr>
<tr>
<td>Mean blood pressure (systolic/diastolic) Method D</td>
<td>212</td>
<td>152.3/84.9 mm Hg (SD 20.7/9.8)</td>
</tr>
</tbody>
</table>

BP = Blood pressure; Method A: first reading; Method B: mean of first two readings; Method C: at least four readings and the mean of the last three readings with less than 15% coefficient of variation difference; Method D: mean of first two consecutive readings with a maximum of 5 mm Hg difference.

19 cases were not recorded in case-record file (CRF).
4 cases were not recorded in CRF.
1 case was not recorded in CRF, and in 10 cases readings were impossible due to pain when readings repeated; large variety in blood pressure readings due to irregular pulse; or sometimes blood pressure over the maximal capacity [>250 mm Hg systolic].
In 8 cases it was not recorded in CRF which reading was first.
4 patients did not have 2 readings.
5 patients did not have 4 readings.

Table 2  Patient characteristics and results of the blood pressure measurement according to the different protocols
study design articles are often published in journals that are not readily available worldwide.

Large hypertension trials are the mainstay for determining cut-off values in guidelines for the diagnosis of hypertension and the target values for treatment. Therefore, measuring protocols used in trials and in guidelines should be consistent across the board, allowing easy comparison worldwide, and they should reflect, and be reflected in, clinical practice. Until uniform BP measuring methods are used in trials, we consider a complete and easily available description of the method used in these trials as an essential minimum requirement in order to be able to translate the results of clinical trials into daily practice recommendations. Only then will the consistency between trials, guidelines and clinical practice improve. The present international effort is directed towards stricter cut-off points and treatment goals. Nevertheless, such goals are easily undermined when professionals cannot agree upon the proper assessment and interpretation of the central point: the measurement method itself.

These results demonstrate that large differences in BP are found when different measuring methods are used. These differences are clinically relevant and show gaps between different methods in trials, guidelines and daily practice. Our study involved only patients with

### Table 3
Differences in blood pressure between the different blood pressure measuring methods

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>215</td>
<td>218</td>
<td>214</td>
<td>208</td>
</tr>
<tr>
<td>MD</td>
<td>0.3</td>
<td>1.6</td>
<td>7.3</td>
<td>7.9</td>
</tr>
<tr>
<td>95% CI</td>
<td>(–0.4 to 1.0)</td>
<td>95% CI (0.6–2.6)</td>
<td>95% CI (5.7–8.9)</td>
<td>95% CI (6.0–9.8)</td>
</tr>
<tr>
<td>P5</td>
<td>14.0% (n = 30)</td>
<td>95% CI (2.1–3.7)</td>
<td>95% CI (4.6–6.9)</td>
<td>95% CI (4.7–7.9)</td>
</tr>
<tr>
<td></td>
<td>95% CI (0.6–2.6)</td>
<td>95% CI (0.6–2.6)</td>
<td>95% CI (4.6–6.9)</td>
<td>95% CI (4.7–7.9)</td>
</tr>
<tr>
<td></td>
<td>95% CI (0.6–2.6)</td>
<td>95% CI (4.6–6.9)</td>
<td>95% CI (4.7–7.9)</td>
<td>95% CI (4.7–7.9)</td>
</tr>
</tbody>
</table>

Grey boxes = systolic blood pressure; White boxes = diastolic blood pressure.
Method A: first reading; Method B: mean of first 2 readings; Method C: at least four readings and the mean of the last three readings with >15% coefficient of variation difference; Method D: mean of first 2 consecutive readings with a maximum of 5 mm Hg difference.
MD: mean difference; P5: percentage patients with >5 mm Hg difference between the two methods; CI = confidence interval.

### Table 4
Sensitivity, specificity, predictive values and likelihood ratios for a positive and negative test of different measures in predicting the UKPDS-method blood pressure threshold of $\geq 150/85$ mm Hg

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Predictive value (positive) (%)</th>
<th>Predictive value (negative) (%)</th>
<th>Likelihood ratio (positive)</th>
<th>Likelihood ratio (negative)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method A</td>
<td>95.0 (90.1–97.7)</td>
<td>66.1 (52.9–77.4)</td>
<td>87.9 (81.9–92.2)</td>
<td>83.7 (69.8–92.2)</td>
<td>2.8 (2.0–4.0)</td>
<td>0.08 (0.04–0.15)</td>
</tr>
<tr>
<td>Method B</td>
<td>98.8 (95.1–99.8)</td>
<td>71.0 (57.9–81.4)</td>
<td>89.8 (84.2–93.7)</td>
<td>95.7 (84.0–99.3)</td>
<td>3.4 (2.3–5.0)</td>
<td>0.02 (0.01–0.07)</td>
</tr>
<tr>
<td>Method D</td>
<td>89.4 (83.4–93.5)</td>
<td>85.5 (73.7–92.7)</td>
<td>94.1 (88.8–97.1)</td>
<td>75.7 (63.7–84.8)</td>
<td>6.2 (3.4–11.3)</td>
<td>0.12 (0.08–0.19)</td>
</tr>
</tbody>
</table>

*With 95% CI.
Method A: first reading; Method B: mean of first two readings; Method C: at least four readings and the mean of the last three readings with <15% coefficient of variation difference; Method D: mean of first two consecutive readings with a maximum of 5 mm Hg difference.
FIGURE 1  Bland and Altman plots of differences between UKPDS method (Method 3) and the other methods. The dotted horizontal line represents the mean difference between each pair of measurements, the broken line represents the regression line. **Correlation is significant at the 0.01 level (2-tailed); *Correlation is significant at the 0.05 level (2-tailed). Method A: first reading; Method B: mean of first 2 readings; Method C: at least four readings and the mean of the last three readings with less than 15% coefficient of variation difference; Method D: mean of first 2 consecutive readings with a maximum of 5 mm Hg difference.
diabetes, so it is not certain that the same conclusions may be applied to the non-diabetic population.

Acknowledgements

We would like to thank Professor E. O’Brien, director of the Blood Pressure Unit at Beaumont Hospital in Dublin, for critically reading a previous version of this manuscript, and Thea Schenk for translating this manuscript.

Conflict of interest statement: none

This work was funded by the Medical Research Fund Zwolle and by the Langerhans Foundation.

References

30 Keane WF, Brenner BM, de Zeeuw D et al. The risk of developing end-stage renal disease in patients with type 2 diabetes and...
34 The HOPE study investigators. The HOPE (Heart Outcomes Prevention Evaluation) Study: the design of a large, simple randomized trial of an angiotensin-converting enzyme inhibitor (ramipril) and vitamin E in patients at high risk of cardiovascular events. *Can J Cardiol* 1996; 12: 127–137.