Letter to the Editor

The 'International Protocol': more insight or more arithmetic?

Richard L. Braam, Theo Thien Department of Internal Medicine, Division of Hypertension and Vascular Pathology, University Medical Centre St. Radboud, Nijmegen, The Netherlands.

It was with interest that we read the recent article concerning the 'International Protocol'—a newly proposed protocol for the validation of automated blood pressure measuring devices [1]. The objective of this new protocol was to simplify the most widely used protocols currently available, namely the British Hypertension Society (BHS) protocol 1993 and the protocol of the Association for the Advancement of Medical Instrumentation (AAMI). This has principally been achieved by eliminating phases 1 to 3 of the BHS-protocol, by decreasing the required number of subjects from 85 to 33 and by relaxing the recruitment of subjects in high and low blood pressure ranges. The 'International Protocol' is two-phased, making it possible to eliminate hopeless devices at an early stage.

Despite these obvious improvements we would like to make some critical remarks. In the 'International Protocol' the A, B, C, D grading system has been replaced by a pass/fail system. This hinders direct comparison between validated devices. In our opinion the best measures to describe a device's performance are the mean of differences and the standard deviation of differences (SDD). However these measures are not advocated by the current protocol. We find this strange not in the least because the minimum required number of differences < 5, < 10 and < 15 mmHg (respectively 65, 80 and 95; Table 2b of the 'International Protocol') are originally based on a normal distribution with a mean error of 0 mmHg and a standard deviation of approximately 5 mmHg.

It is also stated that a large mean difference is usually accompanied with a greater standard deviation of differences; i.e. the standard deviation increases with error [1]. Using the same validation studies that formed the basis for the current changes in protocol we found a correlation coefficient of 0.19 for systolic and −0.09 for diastolic blood pressure between mean error and standard deviation, (see Fig. 1) [2–15]. We therefore claim that it is possible for a device to show a large mean error with a relative small SDD. For such a device, simple correction of blood pressure readings with a constant factor would be appropriate [16]. Applying the new 'International Protocol' would probably classify such a device as unsuitable at the early phase. We therefore recommend studying whether, on the basis of earlier validation reports, it is possible to restore and redefine the AAMI criteria.

To test whether a device can accurately determine blood pressure in individuals a tertiary phase is introduced. In at least 22 out of 33 individuals two out of three differences should be < 5 mmHg and a difference > 5 mmHg for all three comparisons is allowed in no more than three subjects. Large blood pressure fluctuations over time in a few individuals could therefore result in the failure of an accurate device to pass. We support the idea of Shirasaki *et al.*, to correct the SDD by subtracting the standard deviation of individual blood pressure variation from the overall standard deviation of differences [17]. This would allow the user to adjust for the influence of intra-subject variability on the calculated accuracy of devices.

Fig. 1
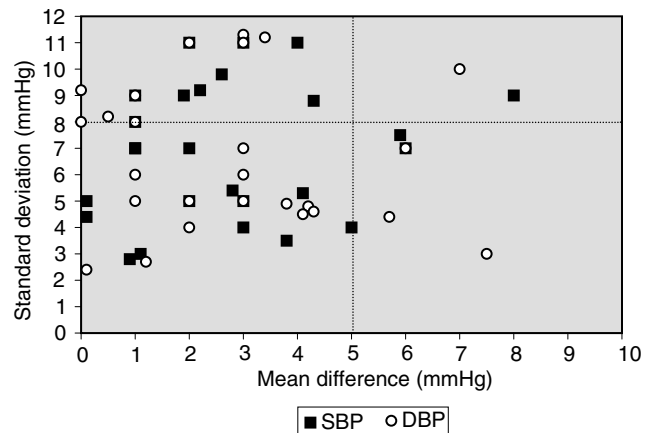
Mean difference versus standard deviation for different blood pressure measuring devices. Data are derived from the validation studies used to adapt the British Hypertension Society protocol [1,3–16]. For some devices more than one combination of mean error and standard deviation is used. For systolic blood pressure (SBP) a correlation coefficient of 0.19 and for diastolic blood pressure (DBP) of −0.09 is found. The vertical and horizontal lines are based on AAMI-criteria.

References

1 O'Brien, Pickering T, Asmar R, Myers M, Parati G, Staessen J, *et al.* on behalf of the Working Group on Blood Pressure Monitoring of the European Society of Hypertension. 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* 2002; **7**:3–17.

2 O'Brien E, Mee F, Atkins N, O'Malley K. Inaccuracy of the Hawksley random-zero sphygmomanometer. *Lancet* 1990; **336**:1465–1468.

3 O'Brien E, Mee F, Atkins N, O'Malley K. Inaccuracy of seven popular sphygmomanometers for home-measurement of blood pressure. *J Hypertens* 1990; **8**:621–634.

4 O'Brien E, O'Ferrall J, Galvin J, Costelloe R, Lydon S, Sheridan J *et al.*, Evaluation of the Accutracker II non-invasive ambulatory blood pressure recorder according to the AAMI Standard. *J Ambul Monitor* 1991; **4**:27–33.

5 O'Brien E, Mee F, Atkins N, O'Malley K. Evaluation of the SpaceLabs 90202 according to the AAMI standard and BHS criteria. *J Hum Hypertens* 1991; **5**:223–226.

6 O'Brien E, Mee F, Atkins N, O'Malley K. Accuracy of the Del Mar Avionics Pressurometer IV determined by the British Hypertension Society Protocol. *J Hypertens* 1991; **9(suppl 5)**:S1–7.

7 O'Brien E, Mee F, Atkins N, O'Malley K. Accuracy of the Novacor DIASYS 200 determined by the British Hypertension Society Protocol. *J Hypertens* 1991; **9(suppl 5)**:S9–15.

8 O'Brien E, Mee F, Atkins N, O'Malley K. Accuracy of the SpaceLabs 90207, Novacor DIASYS 200, Del Mar Avionics Pressurometer IV and Takeda TM-2440 ambulatory systems according to the AAMI and BHS criteria. *J Hypertens* 1991; **9(suppl 6)**:S332–333.

9 O'Brien E, Mee F, Atkins N, O'Malley K. Short report: accuracy of the SpaceLabs 90207 ambulatory blood pressure measuring system in normotensive pregnant women determined by the British Hypertension Society Protocol. *J Hypertens* 1993; **11(suppl 5)**:869–873.

10 O'Brien E, Mee F, Atkins N, O'Malley K. Accuracy of the Dinamap portable monitor, Model 8100, determined by the British Hypertension Society protocol. *J Hypertens* 1993; **11**:761–763.

11 O'Brien E, Mee F, Atkins N, O'Malley K. Accuracy of the CH-Druck/Pressure ERKA ambulatory blood pressure measuring system determined by the British Hypertension Society protocol. *J Hypertens* 1993; **11(suppl 2)**:1–7.

12 O'Brien E, Mee F, Atkins N, O'Malley K. Accuracy of the Profilomat ambulatory blood pressure measuring system determined by the British Hypertension Society protocol. *J Hypertens* 1993; **11(suppl 2)**:S9–15.

13 O'Brien E, Mee F, Atkins N, Thomas M. Evaluation of three devices for self-measurement of blood pressure: according to the revised British Hypertension Society Protocol: the Omron HEM-705CP, Philips HP5332 and Nissei DS-175. *Blood Press Monit* 1996; **1**: 55–61.

14 Mee F, Atkins N, O'Brien E. Evaluation of the Profilomat II^R ambulatory blood pressure system according to the protocols of the British Hypertension Society and the Association for the Advancement of Medical Instrumentation. *Blood Press Monit* 1999; **3**:353–361.

15 O'Brien E, Mee F, Atkins N. Evaluation of the Schiller BR-102 ambulatory blood pressure system according to the protocols of the British Hypertension Society and the Association for the Advancement of Medical Instrumentation. *Blood Press Monit* 1999; **4**:35–43.

16 Braam RL, Thien T. The influence of the method of data analysis on the reported accuracy of automated blood pressure measuring devices. *J Hypertens* 2001; **19**:1765–1767.

17 Shirasaki O, Tereda H, Niwano K, Nakanishi T, Kanai M, Miyawaki Y, *et al*. The Japan Home-health Apparatus Industrial Association: investigation of home-use electronic sphygmomanometers. *Blood Press Monit* 2001; **6**:303–307.

Reply

Eoin O'Brien and Neil Atkins on behalf of the Working Group on Blood Pressure Monitoring of the European Society of Hypertension.

Doctors Braam and Thien kindly acknowledge that the recently published International Protocol from the Working Group on Blood Pressure Monitoring of the European Society of Hypertension [1] incorporates 'obvious improvements' over the earlier protocols of the British Hypertension Society (BHS) [2] and the Association for the Advancement of Medical Instrumentation (AAMI) [3].

However, they are critical of some aspects of the International Protocol.

Their first concern is that replacement of the A–D grading system by a pass/fail system will hinder direct comparison between validated devices. The purpose of this approach was simply an acknowledgement of previous policy whereby devices gaining A and B grades according to the BHS protocol were recommended for clinical use, whereas those with C and D grades were not recommended [2]. Moreover, the tables in the International Protocol give full details of how the results are calculated and provide a more comprehensive means of device comparison than the former grading system.

The use of the mean and standard deviation as the basis for assessing the performance of a device is recommended in the AAMI protocol [3]. However, this approach is founded on the false assumption that device errors are normally distributed around the mean error. If the International Protocol had used a distribution based on a mean error of 0 mmHg and a standard deviation of 5 mmHg, the requirements for < 5 mmHg, < 10 mmHg and < 15 mmHg would have been 67, 94 and 99 measurements respectively. It is not mathematically possible to choose a simple mean and standard deviation that will give a distribution comparable to that in the protocol. If, for example, the AAMI limits of a mean error of 5 mmHg and a standard deviation of 8 mmHg were used, the < 5 mmHg, < 10 mmHg and < 15 mmHg requirements would be 39, 70 and 89 measurements. Even relating the standard deviation to the mean does not help. If, for example, the standard deviation were set so that at most 85% of the measurements would have an error of 10 mmHg, then depending on the mean difference, there would be an expected 5 mmHg limit of between 48 and 53% of the measurements and a restrictive 15 mmHg limit of between 97 and 98%. If values are chosen to ensure that the percentage of accurate measurements at a particular limit are reasonable then the requirements at lower limits will be too liberal whereas those at higher limits will be too restrictive. The use of non-parametric limits in the International Protocol is a valid, simple, and meaningful solution to this problem.

The International Protocol does not state that 'a large mean difference is usually accompanied with a greater standard deviation of differences'; what it does say is that 'standard deviations tend to increase with the error', which is quite different and indeed is supported by the data of Braam and Thien. It would not be appropriate to pass a device with a large mean error with a relatively small SDD as long as a simple constant correction factor was provided, because it would be totally impractical for manufacturers to sell devices and expect users to employ a correction factor. The onus should be on the manufacturer to do this prior to

submitting the device for validation. The purpose of Phase 1 in the International Protocol is to detect such devices at an early stage so as not to dissipate resources on proceeding with a validation that is doomed to failure.

The last issue of concern relates to the tertiary phase of the International Protocol in which a 'difference of 5 mmHg for all three comparisons is allowed in at most three subjects'. This has been introduced to allow specifically for the 'large blood pressure fluctuations over time in a few individuals'. Based on the evidence of previous validation studies good devices will meet this criterion, and devices that do not are inaccurate by definition. It is accepted that studies based on statistical analyses can have Type I and Type II errors, and that a small percentage of devices (mostly marginal ones) will either incorrectly pass or fail a particular validation study. However, it is hoped that the much-simplified International Protocol will result in the same devices being validated in a number of different centres thus reducing greatly the probability of such errors.

Finally, the overall concern that the International Protocol may fail 'accurate' devices that have a 'few' shortcomings, may be countered by the argument that poor devices are being recommended on the basis of the current validation criteria being applied in the AAMI protocol [3]. We are grateful to Braam and Thien for allowing us this opportunity to clarify these important aspects of the International Protocol.

References

1 O'Brien E, Pickering T, Asmar R, Myers M, Parati G, Staessen J, *et al.* on behalf of the Working Group on Blood Pressure Monitoring of the European Society of Hypertension. 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* 2002; **7**:3–17.

2 O'Brien E, Petrie J, Littler WA, de Swiet M, Padfield PL, Altman D, *et al.* The British Hypertension Society Protocol for the evaluation of blood pressure measuring devices. *J Hypertens* 1993; **11(Suppl 2)**:S43–S63.

3 American National Standard. *Electronic or automated sphygmomanometers.* Arlington, VA: Association for the Advancement of Medical Instrumentation; 1993.