Impact of applying the more stringent validation criteria of the revised European Society of Hypertension International Protocol 2010 on earlier validation studies
George S. Stergiou a, Nikos Karpettas a, Neil Atkins b and Eoin O’Brien c

Objective Since 2002 when the European Society of Hypertension International Protocol (ESH-IP) was published it has become the preferred protocol for validating blood pressure monitors worldwide. In 2010, a revised version of the ESH-IP with more stringent criteria was published. This study assesses the impact of applying the revised ESH-IP criteria.

Methods A systematic literature review of ESH-IP studies reported between 2002 and 2010 was conducted. The impact of applying the ESH-IP 2010 criteria retrospectively on the data reported in these studies was investigated. The performance of the oscillometric devices in the last decade was also investigated on the basis of the ESH-IP criteria.

Results Among 119 published studies, 112 with sufficient data were analyzed. According to ESH-IP 2002, the test device failed in 19 studies, whereas by applying the ESH-IP 2010 criteria in 28 additional studies increased the failure rate from 17 to 42%. Of these 28 studies, in 20 (71%) the test device failed at part 1 (accuracy per measurement) and in 22 (79%) at part 2 (accuracy per subject). Most of the failures involved the ‘5 mmHg or less’ criterion. In the last decade there has been a consistent trend toward improved performance of oscillometric devices assessed on the basis of the ESH-IP criteria.

Conclusion This retrospective analysis shows that the stricter revised ESH-IP 2010 criteria will noticeably increase the failure rate of devices being validated. Oscillometric devices are becoming more accurate, and the revised ESH-IP by acknowledging this trend will allow more accurate devices to enter the market.


Keywords: blood pressure, blood pressure monitor, European Society of Hypertension International Protocol, validation

Introduction Electronic blood pressure (BP) monitors are being used increasingly for ambulatory, home, and also office/clinic BP measurement worldwide. Protocols to test the accuracy of these devices have been developed first in 1987 by the American Association for the Advancement of Medical Instrumentation (AAMI) [1] and in 1990 by the British Hypertension Society (BHS) [2]. Revisions of both these protocols were published in 1993 [3,4] and the AAMI protocol was revised again in 2003 [5] and 2009 [6]. Although there are differences between these two protocols, they both standardized the validation procedure and established criteria of accuracy, which allow direct comparison of the performance among different validation studies [1–4].

On the basis of the experience of validation studies the European Society of Hypertension (ESH) working group on BP monitoring developed the ‘ESH-International Protocol’ (ESH-IP) in 2001 [7]. The purpose of this protocol was to simplify the validation procedure and to reduce the sample size required without losing the evaluation accuracy of the earlier more complicated, cumbersome, and costly protocols [1–4]. A recent review showed that from the publication of the ESH-IP in 2002 until June 2010, 48 studies have been reported using the BHS protocol, 38 using the AAMI standard, and 104 using the ESH-IP [8]. Thus, it seems that the ESH-IP that succeeded in expanding by three to four-fold the use of validation procedures worldwide compared with the period before its publication [8].

In 2010, a revised version of the ESH-IP protocol was published [9]. There are several changes in the revised protocol, including (i) relaxing of the lower-age criterion for inclusion and the BP limits in the high range under certain conditions, (ii) requirements for the distribution of observer measurements, (iii) removal of phase 1 analysis, (iv) standardization of reporting the validation results, and, most importantly, (v) tightening of the validation criteria for the pass level. The latter is clearly the most challenging change and seems to be timely, given that there is evidence that developments in technology of BP monitors allow most of the new devices to pass the
validation criteria of the ESH-IP 2002. Indeed, in the above-mentioned review, of 78 validation studies carried out using the ESH-IP, 66 have passed (84.6%) and only 12 (15.4%) failed [8].

The objective of this study was to assess the impact of applying the more stringent validation criteria of the revised ESH-IP 2010 on published validation studies in which the tested devices have passed the ESH-IP 2002. An additional objective was to investigate whether the accuracy of oscillometric devices has improved in the last decade as judged by published validation study data.

Methods
A systematic literature review was conducted for studies in a full paper or an abstract form reporting the results of validations of BP-measuring devices according to the ESH-IP 2002 from its publication in February 2002 to the end of January 2010. Sources were PubMed, abstract books of main meetings of the European, the International and the American Society of Hypertension from 2003 to 2009, and the www.dableducational.org website. Keywords used were: BP measurement, BP devices, validation, European Society of Hypertension International Protocol (ESH-IP). When both a full paper and an abstract were found from the same validation study (same authors and device) the data from the full paper were used.

The impact of the revised ESH-IP 2010 was assessed by retrospectively applying the new more stringent criteria of part 1 and part 2 of the protocol (which have replaced the phase 2.1 and phase 2.2 of the ESH-IP 2002) to the results reported in the earlier validation studies. To investigate whether the accuracy of the oscillometric devices has improved since 2002, the validation studies were grouped according to the year of publication and trends in the validation results in terms of the reported observer–device BP differences evaluated (within 5, 10, and 15 mmHg differences and standard deviation of the mean differences).

Results
A total of 119 validation studies [10–93] carried out using the ESH-IP 2002 and involving 97 devices were identified (Fig. 1). Seven studies in three papers were excluded [10–13] because of insufficient data reported for phase 2.1 and/or phase 2.2 of the ESH-IP (all reported to pass the ESH-IP 2002). The revised ESH-IP 2010 criteria were applied in 112 validations (three in an abstract form only [14,15] and 109 published in a full paper). The characteristics of the devices tested in these 112 validation studies (oscillometric vs. auscultatory devices; arm vs. wrist devices, office vs. home vs. ambulatory use) are presented in Table 1. Information was lacking on the measurement method (oscillometric or auscultatory) in three studies [15–17], on the device design (arm or wrist device) in one study [15], and on the use of the device (office, home, or ambulatory use) in five studies [15,18,19]. All validation studies were treated independently though, in some instances, the same device model was validated in more than one study.

Of the 112 studies evaluated using the ESH-IP 2002, the test device failed in 19 studies [14,18–30], whereas when the more stringent ESH-IP 2010 criteria were retrospectively applied it failed in 28 more studies (Table 2) [18,21,29,31–53]. Thus, using the revised ESH-IP 2002 criteria the protocol fail rate was increased from 17% (19 of 112 studies) to 42% (47 of 112 studies; Fig. 1). The numbers of validation studies that failed the ESH-IP 2002 or the revised ESH-IP 2010 according to the type of the device are presented in Table 1.

Seven of the devices presented in Table 2 have also been validated using the AAMI and/or BHS protocols one using the AAMI protocol [45], two using the BHS [38,44], and four using both AAMI and BHS protocols [21,33,37,49], and all these validations have been successful.

Of the 28 validation studies that passed the ESH-IP 2002 but did not fulfill the criteria of the ESH-IP 2010, 20 (71%) failed at part 1 [18,21,29,33,34,36–44,47,48,52,53]
and 22 (79%) at part 2 [18,21,29,31–33,37,39–41,43,45–53]. These parts correspond to phases 2.1 (testing the device accuracy per BP measurement) and 2.2 of the ESH-IP 2002 (testing the accuracy per subject) (Table 3).

Six devices failed at part 1 [34,36,38,42–44], eight at part 2 [30,31,33,37,39–41,43,47,48,52,53]. At both parts 1 and 2, most of the failures involved systolic BP, some involved diastolic BP, and a few involved both. Considering part 1, the ‘at least two of’ criterion seemed to be more stringent than the ‘all of’ criterion (16 studies [18,29,34,36–38,40–44,47,48,53] vs. 11 [18,21,29,33,34,36,39,40,43,47,52] studies’ failures, respectively) (Table 3).

To determine whether there had been any improvement in the accuracy of the oscillometric devices tested using the ESH-IP 2002 in the last 8 years, 87 of 93 validation studies that provided a complete report of phase 2.1 were assessed. As shown in Fig. 2, in the last decade there was

Table 2 Devices that passed the European Society of Hypertension-International Protocol 2002 but seem to fail when the revised 2010 criteria are retrospectively applied

<table>
<thead>
<tr>
<th>Manufacturer</th>
<th>Model</th>
<th>Arm/wrist</th>
<th>Comments</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>A&amp;D</td>
<td>UA-85X</td>
<td>Arm</td>
<td></td>
<td>[50]</td>
</tr>
<tr>
<td>Andon</td>
<td>KD-391</td>
<td>Arm</td>
<td></td>
<td>[47]</td>
</tr>
<tr>
<td>Braun</td>
<td>BP 3550</td>
<td>Wrist</td>
<td></td>
<td>[40]</td>
</tr>
<tr>
<td>Datascpe</td>
<td>Accutor Plus</td>
<td>Arm</td>
<td>Children</td>
<td>[21]</td>
</tr>
<tr>
<td>Health &amp; Life</td>
<td>HLB66BA</td>
<td>Arm</td>
<td></td>
<td>[48]</td>
</tr>
<tr>
<td>HealthSTATS</td>
<td>BPro</td>
<td>Wrist</td>
<td></td>
<td>[45]</td>
</tr>
<tr>
<td>Microlife</td>
<td>WatchBP O3</td>
<td>Arm</td>
<td>Two independent studies</td>
<td>[46,53]</td>
</tr>
<tr>
<td>Microlife</td>
<td>WatchBP Office</td>
<td>Arm</td>
<td></td>
<td>[31]</td>
</tr>
<tr>
<td>Omron</td>
<td>637IT</td>
<td>Wrist</td>
<td>Three independent studies (adults, obese, elderly)</td>
<td>[41,42]</td>
</tr>
<tr>
<td>Omron</td>
<td>705IT</td>
<td>Arm</td>
<td>Children</td>
<td>[33]</td>
</tr>
<tr>
<td>Omron</td>
<td>Elite 7300W</td>
<td>Arm</td>
<td>Only women</td>
<td>[29]</td>
</tr>
<tr>
<td>Omron</td>
<td>HEM-711 DLX</td>
<td>Arm</td>
<td>Wide arm range (22–42 cm)</td>
<td>[51]</td>
</tr>
<tr>
<td>Omron</td>
<td>HEM-780REL</td>
<td>Arm</td>
<td>Wide arm range (22.8–43.2 cm)</td>
<td>[34]</td>
</tr>
<tr>
<td>Omron</td>
<td>M6 (HEM-7001-E)</td>
<td>Arm</td>
<td>Elderly</td>
<td>[35]</td>
</tr>
<tr>
<td>Omron</td>
<td>M6 Comfort (HEM-7000-E)</td>
<td>Arm</td>
<td></td>
<td>[38]</td>
</tr>
<tr>
<td>Omron</td>
<td>M7 (HEM-780E)</td>
<td>Arm</td>
<td>Only large cuff</td>
<td>[37]</td>
</tr>
<tr>
<td>Panasonic</td>
<td>EW3109</td>
<td>Arm</td>
<td></td>
<td>[18]</td>
</tr>
<tr>
<td>Rossmax</td>
<td>ME 701 series</td>
<td>Arm</td>
<td></td>
<td>[38]</td>
</tr>
<tr>
<td>Schiller</td>
<td>BR-102 plus</td>
<td>Arm</td>
<td>Two independent studies (auscultatory, oscillometric)</td>
<td>[43]</td>
</tr>
<tr>
<td>Sensacare</td>
<td>SAW-102</td>
<td>Wrist</td>
<td></td>
<td>[49]</td>
</tr>
<tr>
<td>SunTech Medical</td>
<td>OSCAR 2</td>
<td>Arm</td>
<td></td>
<td>[44]</td>
</tr>
<tr>
<td>Visomat</td>
<td>Comfort 20/40</td>
<td>Arm</td>
<td></td>
<td>[39]</td>
</tr>
<tr>
<td>Visomat</td>
<td>Comfort form</td>
<td>Arm</td>
<td>Wide arm range (22–42 cm)</td>
<td>[52]</td>
</tr>
</tbody>
</table>

*These results are derived from a retrospective application of validation results and should not be interpreted as actual validation results.

Table 3 Analysis of validation studies that passed the ESH-IP 2002 but fail in the revised 2010 ESH-IP criteria (n=28) presenting the number of devices failing per validation criterion (proportions of failing devices per criterion in parentheses (%); thresholds for ESH-IP criteria 2002 [7] and 2010 [9] in brackets [2002; 2010])

<table>
<thead>
<tr>
<th>ESH-IP 2010 part 1 criteria (measurement accuracy)</th>
<th>≤ 5 mmHg [%]</th>
<th>≤ 10 mmHg [%]</th>
<th>≤ 15 mmHg [%]</th>
<th>‘At least two of’ criterion [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total ‘All of’ criterion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SBP</td>
<td>17 (61)</td>
<td>8 (28)</td>
<td>5 (17)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>DBP</td>
<td>8 (29)</td>
<td>5 (17)</td>
<td>5 (17)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>SBP and DBP</td>
<td>5 (18)</td>
<td>2 (7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SBP and/or DBP</td>
<td>20 (71)</td>
<td>11 (39)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ESH-IP 2010 part 2 criteria (patient accuracy)*</th>
<th>≤ 5 mmHg [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2/3 (≤ 5 mmHg)</td>
<td>[2002: 22]</td>
</tr>
<tr>
<td></td>
<td>[2010: 23]</td>
</tr>
<tr>
<td>SBP</td>
<td>15 (54)</td>
</tr>
<tr>
<td>DBP</td>
<td>10 (34)</td>
</tr>
<tr>
<td>SBP and DBP</td>
<td>3 (10)</td>
</tr>
<tr>
<td>SBP and/or DBP</td>
<td>22 (79)</td>
</tr>
</tbody>
</table>

DBP, diastolic blood pressure; ESH-IP, European Society of Hypertension-International Protocol; SBP, systolic blood pressure.

*The 0/3 ≤ 5 mmHg criterion did not change in the revised protocol.
a consistent trend toward an improved accuracy of the oscillometric devices as indicated by an increased number of device–observer BP differences within 5, 10, or 15 mmHg (Fig. 2, panels a, b, and c, respectively) and a decreased standard deviation of the differences (Fig. 2, panel d).

Discussion
This study assesses the impact of the more stringent criteria of the revised ESH-IP 2010 on earlier studies that have fulfilled the requirements and passed the original ESH-IP 2002. This retrospective analysis showed that approximately one-third of validations that passed the ESH-IP 2002 might not satisfy the criteria of the revised ESH-IP 2010. Thus, if the devices validated according to the ESH-IP were to be validated according to the revision of the ESH-IP, the failure rate would probably increase from 17 to 42%. It should be noted that these results are derived from a retrospective application of validation results and should not be interpreted as actual validation results.

Indeed, one of the main purposes for revising the ESH-IP was to make its requirements more stringent. This seemed to be timely given that the vast majority of devices tested on the basis of the ESH-IP 2002 have been successful [8]. Although a publication bias cannot be excluded (some negative studies not being published), these data suggest that the current technology of BP monitors has improved, thereby allowing an increase in the level of minimal accuracy requirements for approval. Indeed, both parts 1 and 2 of the revised ESH-IP 2010 had a similar impact on the pass by fail rate of the devices. Seven of the devices that passed the ESH-IP 2002 but failed the ESH-IP 2010 (Table 2) have been also validated using the BHS and/or the AAMI protocol and they have all succeeded [21,33,37,38,44,45,49]. This suggests that the ESH-IP 2010 criteria are more stringent than those of the BHS and the AAMI criteria.

A promising feature of this analysis, which is in agreement with the high rate of successful validations in the last decade, is a trend toward an improvement in the accuracy of oscillometric devices in the period between 2002 (when the ESH-IP was first published) and 2010. This is clearly shown by the improvement in the performance of the oscillometric devices when assessed on the basis of several ESH-IP 2002 validation criteria shown in Fig. 2.

In conclusion, this analysis suggests that in the last decade there has been a trend toward an improvement in the performance or accuracy of the oscillometric devices,
which justifies the adoption of more stringent validation criteria. This issue has been successfully addressed in the 2010 revision of the ESH-IP that set higher standards for device validation. The application of the more stringent criteria of the revised ESH-IP in future studies is expected to increase by more than two-fold the validation fail rate.

Acknowledgements
Sources of funding: none declared.

Conflict of interest: GS Stergiou has received consultation fees and research support from Microlife manufacturer of BP monitors.

References
36 Belgazi J, El Feghali RN, Moussalem T, Rejdych M, Asmar RG. Validation of four automatic devices for self-measurement of blood pressure according to


