Accuracy of monitors used for blood pressure checks in English retail pharmacies: a cross-sectional observational study

INTRODUCTION
High blood pressure (BP) is a key risk factor for the development of cardiovascular disease\(^1\) and a major cause of morbidity and mortality worldwide.\(^2\) An accurate BP monitoring device is fundamental to BP measurement in the diagnosis and control of hypertension.

Several protocols\(^3\)–\(^5\) exist for the validation of BP measuring devices but these are generally undertaken and then published on brand new models, and so do not guarantee sustained accuracy thereafter. The revised General Medical Services contract for UK GPs (2003)\(^6\) includes a recommendation to ensure medical equipment is regularly maintained, calibrated and replaced if necessary. Typically, new monitors are assumed to be accurate for 2 years and then annual checks are undertaken. However, it is not clear whether this is appropriate, as the drift in accuracy over time of an automated sphygmomanometer is not known. The error rate is a function of random (variability) and systematic (bias) error, and ultimately depends on the calibration interval of the device, and the conditions under which it is used.

Detection and control of hypertension are sub-optimal,\(^7\) and a recent study has identified that there are insufficient GPs in England to achieve high levels of detection while at the same time maintaining access to appointments with GPs.\(^8\) Community pharmacies are a good potential site for identifying cardiovascular risk factors and improving disease detection,\(^9\) including identifying people with hypertension, because of their accessibility and because many pharmacies provide free access to BP monitors. Pharmacists have been involved in successful community-based screening programmes developed to improve detection and treatment of hypertension, both in the UK and worldwide.\(^10\)–\(^14\) Community pharmacy BP monitoring is readily available, widely demanded, and recommended by hypertension guidelines.\(^11,17\)–\(^18\)

Evidence from the US\(^19\) suggests that many people with hypertension check their BP at a pharmacy. In the UK, the ‘Know Your Numbers’ campaign has ensured that 1.5 million people have had their BP checked in the UK since 2001, mostly at community pharmacies.\(^20\)

However, the few studies evaluating publicly available monitors have shown them to be inaccurate.\(^21\)–\(^27\) Many of these studies are more than a decade old, and there remains limited evidence of how well the monitors maintain their rigour in a potentially challenging environment.

Therefore, the authors assessed the accuracy of validated\(^28\) automatic BP

---

Abstract
Background
Free blood pressure (BP) checks offered by community pharmacies provide a potentially useful opportunity to diagnose and/or manage hypertension, but the accuracy of the sphygmomanometers in use is currently unknown.

Aim
To assess the accuracy of validated automatic BP monitors used for BP checks in a UK retail pharmacy chain.

Design and setting
Cross-sectional, observational study in 52 pharmacies from one chain in a range of locations (inner city, suburban, and rural) in central England.

Method
Monitor accuracy was compared with a calibrated reference device (Omron PA-350), at 50 mmHg intervals across the range 0–300 mmHg (static pressure test), with a difference from the reference monitor of +/-3 mmHg at any interval considered a failure. The results were analysed by usage rates and length of time in service.

Results
Of 61 BP monitors tested, eight (13%) monitors failed (that is, were +/-3 mmHg from reference), all of which underestimated BP. Monitor failure rate from the reference monitor of +/-3 mmHg at any testing interval varied by length of time in use (2/36, 5% <18 months; 4/14, 29% >18 months, \(P = 0.038\)) and to some extent, but non-significantly, by usage rates (4/22, 18% in monitors used more than once daily; 2/33, 6% in those used less frequently, \(P = 0.204\)).

Conclusion
BP monitors within a pharmacy setting fail at similar rates to those in general practice. Annual calibration checks for blood pressure monitors are needed, even for new monitors, as these data indicate declining performance from 18 months onwards.

Keywords
blood pressure monitors; calibration; community pharmacy services; hypertension; primary health care.
METHOD

The accuracy of all available digital sphygmomanometers (two models, both derived from the validated Kinetik BPM 1 Series) in 52 pharmacies was evaluated through comparison to a calibrated reference monitor (Omron PA-350), at 50 mmHg intervals across the range 0–300 mmHg, as recommended by the British Hypertension Society. Following manufacturers’ protocol, a difference from the reference monitor of +/− 3 mmHg at any interval was considered a failure. In addition to this static pressure test, the authors also conducted a visual inspection to check that the machine switched on and off, contained batteries, and had a readable display, and they tested deflation, air leakage, and all cuffs in use. Pharmacies were visited in a range of locations (inner city, suburban, and rural) in central England (Birmingham, Black Country, Herefordshire, and Worcestershire) in order to achieve a sample of monitors of both high and low use. No pharmacies refused to participate. All testing was undertaken by one author.

The length of time in use and number of recorded uses was measured in two ways:

- Pharmacy staff were asked at the time of the calibration visit. Very occasionally it was calculated by reference to a log book, and some monitors had stickers that detailed the precise date they were first used for BP checks, but mainly it was dependent on staff recall. Data on service duration, or length of time in use, were recorded only from the individual pharmacies.
- Figures detailing the number of BP checks for each pharmacy were collected from the head office for the most recent 2-year period; however, this data were at the pharmacy level and not available for individual monitors, and thus were of limited use in pharmacies that had multiple monitors. The data were also dependent on individual data being inputted every time a BP check was performed.

The association between monitor precision and both length of time in service and usage rate (using the data collected from fieldwork and those from the pharmacy head office, separately) was tested using linear regression. The linear models were fitted by using the mean of the differences as outcome, which was defined as the mean of the differences between the device being tested and the reference standard in all testing intervals ascending and descending. All model assumptions were checked. Failure rate by the different predictors was assessed using the Fisher exact test statistic.

RESULTS

Of a total of 61 monitors assessed, eight (13%) failed (that is, were >3 mmHg from reference), all of which underestimated BP. The largest disparity found was 8.3 mmHg at a pressure of 300 mmHg. The majority of failures were at higher pressures, and if a monitor failed at a given pressure it invariably failed at all higher pressures; eight failed at 300 mmHg, five failed from 250 mmHg, four from 200 mmHg, and one from 100 mmHg. At the testing point nearest to the diagnostic threshold (150 mmHg), two monitors failed (2/61, 3%).

Overall, the difference in BP between monitors and the reference device increased with level of BP in a linear fashion to a mean 2 mmHg difference at 300 mmHg (Table 1). The relative difference did not appear to vary with pressure, as the percentage error was almost identical at each testing interval, albeit a little higher at the point closest to the diagnostic threshold in the testing process (150 mmHg), where there was a mean difference of just over 1 mmHg (Table 1).

BP was underestimated in 71% of readings and overestimated in 22%, and the ratio of underestimated readings to overestimated readings increased with pressure: 80%
of readings were underestimated and 16% were overestimated between 150–250 mmHg, with 92% underestimated and 5% overestimated at 300 mmHg (Table 2). This pattern of underestimation is suggestive of a small systematic error across the monitors, but with random error (variability) in individual monitors.

Estimates from pharmacy staff of annual usage of BP monitors ranged from 104 times to 1560 times (compounded by variation from the ‘Know Your Numbers’ campaign). The pharmacy chain’s head office range of estimated annual usage was from 14–276 times.

Comparison of the ratio of uses between the two data sources was inconsistent, varying from a single pharmacy where usage data from head office was higher, to monitoring usage from fieldwork estimates being >18 times greater. In view of this, both estimates were used separately and compared in sensitivity analyses.

Length of time in service varied from 1 week to almost 3 years, with a modal period of service duration of 6–12 months. However, there was some uncertainty about the time in use of several monitors, with no clear estimate available for nine (15%) of the devices.

Monitor failure rate (defined as a difference from the reference monitor of +/– 3 mmHg at any testing interval) varied by length of time in use (2/38, 5% <18 months compared with 8/61, 13% overall and versus 4/14, 29% >18 months, P = 0.038). To some extent failure rate varied by usage rates (4/22, 18% in monitors used more than once daily compared with 2/33, 6% in those used less frequently) but this difference was not statistically significant (P = 0.204). A greater failure rate (4/13, 31%) was observed for higher usage when the head office data were used, but as this was only available at pharmacy level, and four failures happened in pharmacies where multiple devices were being used, the data were difficult to interpret.

The association between monitor precision and length of time in service (Table 3) using linear regression showed that the mean difference for monitors in service for both 6–12 months and 13–18 months was around 0.35 mmHg greater than that for those used for 0–5 months, increasing to >0.70 mmHg for monitors used for 19–24 months. However, no statistical significance could be observed for any of the categories.

A linear trend for predicted difference was found by amount of usage (for both the fieldwork data and that from the pharmacy’s head office), suggesting a drop in precision with high usage, but the results were not

### Table 1. Mean difference against reference standard at all testing intervals

<table>
<thead>
<tr>
<th>Pressure range, mmHg</th>
<th>Mean difference, mmHg (95% CI)</th>
<th>Error of the mean difference, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0 (–0.01 to 0.01)</td>
<td>–</td>
</tr>
<tr>
<td>50</td>
<td>0.15 (0.01 to 0.30)</td>
<td>0.30</td>
</tr>
<tr>
<td>100</td>
<td>0.42 (0.21 to 0.62)</td>
<td>0.42</td>
</tr>
<tr>
<td>150</td>
<td>1.15 (0.32 to 1.98)</td>
<td>0.76</td>
</tr>
<tr>
<td>200</td>
<td>1.14 (0.84 to 1.44)</td>
<td>0.57</td>
</tr>
<tr>
<td>250</td>
<td>1.41 (1.04 to 1.78)</td>
<td>0.56</td>
</tr>
<tr>
<td>300</td>
<td>1.90 (1.48 to 2.33)</td>
<td>0.63</td>
</tr>
</tbody>
</table>

Pressure is checked in 50 mmHg increments from 0–300 mmHg and in 50 mmHg increments down again. The data at each testing interval between 0–250 mmHg going up have been combined with the data going down.

### Table 2. Overestimation and underestimation of pressure at all testing intervals

<table>
<thead>
<tr>
<th>Pressure range, mmHg</th>
<th>Underestimate, %</th>
<th>Overestimate, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>47</td>
<td>32</td>
</tr>
<tr>
<td>50</td>
<td>58</td>
<td>37</td>
</tr>
<tr>
<td>100</td>
<td>70</td>
<td>25</td>
</tr>
<tr>
<td>150</td>
<td>78</td>
<td>17</td>
</tr>
<tr>
<td>200</td>
<td>81</td>
<td>17</td>
</tr>
<tr>
<td>250</td>
<td>82</td>
<td>14</td>
</tr>
<tr>
<td>300</td>
<td>92</td>
<td>5</td>
</tr>
</tbody>
</table>

Pressure is checked in 50 mmHg increments from 0–300 mmHg, and in 50 mmHg increments down again. The data at each testing interval between 0–250 mmHg going up have been combined with the data going down. Overestimates and underestimates include any difference of 0.1 mmHg or more from the reference device.
Further sensitivity analyses were conducted to check the influence of outliers (monitors with a very high mean difference compared with the others), as some monitors could have been faulty from the outset, but results were consistent with the main analyses. All monitors passed the fast deflation test (deflation from 260 mmHg to 15 mmHg should take <10 seconds) and only one failed the air leakage check (inflating to between 280–290 mmHg and requiring the pressure to drop less than 6 mmHg over 1 minute), but this device also failed the static pressure test.

A total of 52 normal (22–30 cm), 45 large (30–42 cm), and 17 extra-large (42–48 cm) cuffs being used with the monitors were also assessed, all of which passed the fast deflation test and air leakage check.

### DISCUSSION

**Summary**

This study found that 13% of monitors that are used for BP checks in community pharmacies tested by the authors failed a standardised accuracy test, underestimating true BP, with the absolute pressure gap between the machines tested and the reference device increasing as BP rose. There was a trend for an increase in mean difference of monitors with longer service duration and higher usage but not to a statistically significant degree, perhaps due to lack of power. Failure rate of monitors after 18 months in use was significantly higher than for those with shorter service duration. Overall accuracy around the systolic diagnostic threshold of 140 mmHg was good, with only small mean differences from reference.

**Strengths and limitations**

This is the only study that the authors are aware of that considers the accuracy of monitors used in pharmacy BP checks, and one of the very few studies to assess accuracy of monitors with respect to time in service. As monitors in use were limited to two models, it provided an opportunity to consider calibration drift in the community for numbers of monitors not usually available on such a scale.

The weaknesses of the data surrounding length of usage and service duration have been previously noted. Both under- and overestimation appeared dependent on who was asked. Length of time in service varied from 1 week to almost 3 years, which was surprising as the pharmacy chain in question has a policy of replacing the entire stock of monitors used for community testing every 2 years at the same time. The a priori intention in this study was to consider a cohort of monitors at the 18–24-month stage of the cycle, which in theory presented a rare opportunity for a reasonable sample size for this type of research. In reality there was an even spread of service duration. Furthermore, the chain’s policy of jettisoning all monitors at a predetermined point, including replacement monitors (that is, ones that may not have been in use for the full 2-year period), meant the sample size was relatively small, there being limited value in the study continuing further than it did, as all remaining monitors were about to be renewed.

### Table 3. Monitor precision and failure rate by length of time in service

<table>
<thead>
<tr>
<th>Service duration, months</th>
<th>Monitors tested, n</th>
<th>Failure rate, % (n)</th>
<th>Mean difference, mmHg (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–5</td>
<td>10</td>
<td>10 (1)</td>
<td>0.38 (0.19 to 0.95)</td>
</tr>
<tr>
<td>6–12</td>
<td>17</td>
<td>0 (0)</td>
<td>0.74 (0.30 to 1.17)</td>
</tr>
<tr>
<td>13–18</td>
<td>11</td>
<td>9 (1)</td>
<td>0.72 (0.17 to 1.24)</td>
</tr>
<tr>
<td>19–24</td>
<td>9</td>
<td>33 (3)</td>
<td>1.09 (0.49 to 1.69)</td>
</tr>
<tr>
<td>&gt;24</td>
<td>5</td>
<td>20 (11)</td>
<td>1.16 (0.35 to 1.96)</td>
</tr>
<tr>
<td>Unsure*</td>
<td>9</td>
<td>22 (2)</td>
<td>–</td>
</tr>
</tbody>
</table>

*Unsure refers to when staff were unable to provide evidence or estimate the service duration of a monitor.

### Table 4. Monitor precision and failure rate by amount of usage, as reported by pharmacy staff

<table>
<thead>
<tr>
<th>Category of usage</th>
<th>Amount of usage</th>
<th>Monitors tested, n</th>
<th>Failure rate, % (n)</th>
<th>Mean difference, mmHg (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Twice daily or more</td>
<td>12</td>
<td>17 (2)</td>
<td>1.05 (0.55 to 1.55)</td>
</tr>
<tr>
<td>Medium high</td>
<td>More than once daily up to twice daily</td>
<td>10</td>
<td>20 (2)</td>
<td>0.68 (0.14 to 1.22)</td>
</tr>
<tr>
<td>Medium low</td>
<td>More than four times weekly up to once daily</td>
<td>16</td>
<td>6 (1)</td>
<td>0.65 (0.23 to 1.07)</td>
</tr>
<tr>
<td>Low</td>
<td>‘Very little’ — up to four times weekly</td>
<td>17</td>
<td>6 (1)</td>
<td>0.67 (0.19 to 1.15)</td>
</tr>
<tr>
<td>Unsure*</td>
<td>Undefined</td>
<td>6</td>
<td>33 (2)</td>
<td>–</td>
</tr>
</tbody>
</table>

*Unsure refers to when staff were unable to provide evidence or estimate the amount of usage of a monitor.

### Table 5. Monitor precision and failure rate by amount of usage, head office data

<table>
<thead>
<tr>
<th>Category of usage</th>
<th>Annual usage rate</th>
<th>Pharmacies where monitors tested, n</th>
<th>Failure rate, % (n)</th>
<th>Mean difference, mmHg (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>116–275.5</td>
<td>13</td>
<td>31 (4)</td>
<td>1.22 (0.72 to 1.71)</td>
</tr>
<tr>
<td>Medium high</td>
<td>86.5–108.5</td>
<td>10</td>
<td>10 (1)</td>
<td>0.87 (0.34 to 1.41)</td>
</tr>
<tr>
<td>Medium low</td>
<td>60.5–82.5</td>
<td>13</td>
<td>15 (2)</td>
<td>0.78 (0.32 to 1.24)</td>
</tr>
<tr>
<td>Low</td>
<td>13.5–59.5</td>
<td>16</td>
<td>6 (1)</td>
<td>0.47 (0.04 to 0.90)</td>
</tr>
</tbody>
</table>

Data available only at level of pharmacy (four failures happened in pharmacies where multiple devices were being used, making it difficult to interpret the relationship with usage).
As a consequence of the company’s 2-year cycle, no formal calibration checking was carried out but several pharmacy staff anecdotally reported that they rejected any monitors they had concerns about (for example, where a monitor had been dropped or had liquids spilt on it), or alternatively if readings seemed spurious or likely to be wrong, albeit without formal testing against a reference device, sending them back to head office. This is reassuring and acts as a further level of quality control.

**Comparing with existing literature**

The failure rate of 13% identified in this study is very similar to that found by A’Court et al when testing monitors used in GP practices, namely 14% overall and 12% specifically for digital devices. Interestingly, for technical reasons those investigators were not able to test the monitors included in this study, which suggests that the underlying technologies in oscillometric monitors may be similarly robust. Of note, however, A’Court et al found no difference in accuracy by service duration over a 2-year period. Perhaps this discrepancy is due to differing usage rates: the estimates of usage from pharmacy staff correspond to the lower end of BP checks conducted in a typical GP surgery, where a BP machine would generally be used for many individuals per day.

In another study, 26% (12/47) of automated BP monitors in use at a hospital were not within 3 mmHg of a reference device, although only a minority of these machines were validated and it is unclear how long they had been in service. That study also found that calibration errors in automated devices tend to underestimate pressure. Conversely, 5% of automated BP monitors in GP practices failed but these devices were all provided by one manufacturer. There are very few studies that assess the accuracy of BP monitoring equipment over time. One study of 14 identical ambulatory monitors over a 6-year period found that 90% of a standardised set of pressure readings were within 2 mmHg on repeated measurement.

**Implications for practice**

Provision of BP monitoring facilities in community pharmacies has the potential to improve the quality of care for patients but depends on accurate monitors as well as robust communication systems with the rest of the health service. This work supports current recommendations for testing monitor accuracy annually but suggests that leaving monitors for as long as 2 years from purchase to testing may lead to potential for error. The main issues detected in this study were of underestimation of BP. Although pharmacists referred patients back to their GP rather than initiating treatment, reducing safety concerns associated with overdiagnosis, any underdiagnosis due to monitor underestimation of BP around clinical threshold levels represents a missed opportunity of pharmacy screening and management of hypertension, although most of the failures were at pressures higher than this. The effect could be exaggerated as previous research has highlighted community pharmacy-measured BP is lower than clinic-measured BP due to a reduced white-coat effect, even though a similar threshold for hypertension status is used by GPs and pharmacists.

This research has implications for patients who self-monitor their BP, as these type of monitors retail on the high street and are very popular, because they are validated and inexpensive, selling up to millions of units. However, no advice is given (including in the manual provided) as to how long these monitors should be used for or how regularly they should be calibrated. Clearly, the amount of usage monitors will receive when purchased by one person for use in the home environment is less than those used regularly in pharmacy BP checks, but the accuracy of such monitors cannot be assumed indefinitely, especially as at least 30% of patients diagnosed with hypertension possess home BP measuring devices.

Successive governments’ national policy has argued that expanding the range of services pharmacies provide will increase access and patient choice, reduce GP workload, and lower NHS costs. This work shows that one building block — namely monitor accuracy — can be safely put into place. However, the authors found large differences in the uptake of the BP checks in different pharmacies within the same chain, and this highlights the importance of implementation. More could be done to encourage people to use this service or to understand why some facilities are apparently more attractive than others.

Community pharmacy BP checks present an excellent opportunity to improve hypertension diagnosis and management, but require accurate equipment. This study has shown BP monitors within this setting to be of similar accuracy to those in primary care, but that without similar calibration checks there is potential for error. These data indicate an annual calibration check is needed for this type of monitor, with evidence suggesting declining performance from 18 months onwards.

**Funding**

This work represents independent research commissioned by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research funding scheme (RP-PG-0407-10347). The views expressed in this study are those of the authors and not necessarily of the NHS, the NIHR, or the Department of Health. Richard J McManus is supported by an NIHR professorship (NIHR-RP-02-12-015) and by the NIHR Collaboration for Leadership in Applied Health Research and Care (CLAHRC) Oxford at Oxford Health NHS Foundation Trust. FD Richard Hobbs is part-funded as director of the NIHR School for Primary Care Research (SPCR), theme leader of the NIHR Oxford Biomedical Research Centre (BRC), and director of the NIHR CLAHRC Oxford. No funding was received from either the pharmacy chain or the manufacturer of the monitor.

**Ethical approval**

The proposal required approval only from the university ethics board, as it involved testing of machines and not intervention with subjects; thus it was classified as service development not requiring wider ethical approval.

**Provenance**

Freely submitted; externally peer reviewed.

**Competing interests**

Richard J McManus has received research support for blood pressure monitors from Omron and Lloyds Pharmacies, and funding to attend and speak at the Japanese Society of Hypertension.

**Acknowledgements**

The authors would like to acknowledge the help of the pharmacies involved in this research, without whose input it would not have been possible. Also we acknowledge the representatives of Kinetik for providing us with the relevant valves and maintenance manual required to undertake the monitor testing.

**Discuss this article**

Contribute and read comments about this article: [bjgp.org/letters](http://bjgp.org/letters)
REFERENCES


