Assessing agreement of blood pressure-measuring devices in Tibetan areas of China: a systematic review

Cuomu Mingji,1,2 Igho J Onakpoya,3 Carl J Heneghan,3 Alison M Ward3

ABSTRACT
Background The validity of blood pressure (BP)-measuring tools at very high altitudes is uncertain. Therefore, the objective of this review was to examine the degree of agreement of BP-measuring devices in Tibet.

Methods We conducted electronic searches in Medline, Embase, Cinahl, Cochrane Library, Global Health Library and the ISI Web of Science. Randomised and observational studies were considered for inclusion. The methodological characteristics of included studies were assessed using the Quality Assessment of Diagnostic Accuracy Studies-2 criteria. Our primary outcome was the difference in mean BP measurements between the new device and the gold standard.

Results We identified three eligible studies, out of which two with a total of 162 participants were included. The studies differed in their methodology: one study reported significantly higher systolic blood pressure (SBP) measurement with electronic sphygmomanometer (Omron) compared with mercury sphygmomanometer (mean difference 5.8±4.7 mm Hg; p<0.001), with no significant difference in diastolic blood pressure (DBP) measurement (0.4±3.9 mm Hg; p=0.23). The second study reported mean differences of 1.0±5.9 mm Hg and −3.1±4.6 mm Hg for SBP and DBP, respectively.

Conclusion The limited evidence from published studies suggests that automated (Omron) BP monitors show a high degree of agreement for DBP when compared against mercury sphygmomanometer at high altitudes. However, the degree of such agreement for SBP is not consistent. Few studies assessing the validity of automated BP monitors at high altitudes have been conducted, and they differ in design and methodology. Further research assessing the suitability of BP-measuring instruments at high altitudes is therefore warranted.

INTRODUCTION
Hypertension is one of the most preventable risk factors for cardiovascular disease (CVD).1 2 However, while the prevalence of hypertension is decreasing in high-income countries due to healthcare interventions, the prevalence of hypertension remains high in low-income and middle-income countries.3 Risk factors such as diet, alcohol, physical exercise and tobacco are the focus of noncommunicable disease programmes.4 5

Tibet is located on the high Qinghai-Tibetan plateau and is known as the ‘third pole’ of the world. Its average altitude is around 4500 m above sea level. The life expectancy in Tibet is lower than that in the mainland by 9 years.6 Although the cause of the shortened life expectancy is unclear, results of observational studies have suggested that there is a higher prevalence of hypertension in Tibet compared with mainland China.7 8 In addition, there is evidence of a higher mortality rate from CVD-related morbidities (eg, stroke) in Tibet.9 Furthermore, a positive correlation between high altitude and the prevalence of hypertension among Tibetans was shown in our recent review.10 However, the results of several studies examining the association of hypertension and high altitude are not always consistent.11–17 The inconsistencies could be due to variations in the methods used to measure blood pressure (BP) across the studies18 or differences in the effect of altitude on the BP-measuring devices.19

Accurate BP measurements are a cornerstone vital for the diagnosis of hypertension,14 and correct BP recordings are important for the reduction of CVD-related morbidity and mortality in low-resource settings.20 Several BP-measuring devices have been validated for use at normal altitudes15–17 and a number of studies testing the accuracy of such devices at higher altitudes have also been conducted.22 However, the degree of agreement of BP-measuring tools in Tibet has not been systematically reviewed. Therefore, the objective of this review was to examine the degree of agreement of BP-measuring devices in Tibet, using data from published studies.

METHODS
Electronic searches were conducted in the following databases: Cinahl; Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews and Database of Abstracts of Reviews of Effects; Medline; Embase; Global Health Library and the ISI Web of Science. Each database was searched from inception until April 2015. No age, time or language restrictions were imposed. Search terms used included ‘Tibet’, ‘Hypertension’, ‘Blood Pressure Determination’, ‘sphygmomanometer’ and ‘diagnosis’ (see online supplementary appendix 1 for a comprehensive search strategy). We also searched Google Scholar for relevant internet proceedings and searched the bibliographies of retrieved full texts.

Randomised and observational studies were considered for inclusion in the review. Studies not conducted in Central Tibet or Tibetan provinces of China were excluded. Two reviewers (CM and IO) independently determined study eligibility. Because tests for accuracy of new devices rely on agreement between the new device and the reference standard, the methodological quality of included studies was
independently assessed by the two reviewers (CM and IJO) using the revised Quality Assessment of Diagnostic Accuracy Studies-2 criteria, which rate the risk of bias and applicability concerns by examining the following domains: patient selection, index test, reference standard and flow and timing. Disagreements were resolved through consensus. Our primary outcome measure was the difference in mean (raw and adjusted) systolic BP (SBP) and diastolic BP (DBP) measurements between the new device and the gold standard (mercury sphygmomanometer).

Data were extracted according to the study design, the previous validation of devices used in the studies, the qualifications of the observers, the data collection procedures, setting, age, gender, altitudes, the results of the BP readings and the sources of funding. We used frequency graphs to display the methodological ratings of included studies, and summary tables to present the main results of included studies. We had planned to statistically combine the data from included studies.

RESULTS

Our searches identified 196 potential articles, out of which three studies were considered eligible (figure 1). One study was excluded because it was a duplicate article of another study that was included in the review. Finally, two articles with a total of 162 participants were included in the review.

Both studies were observational in design and were conducted in Lhasa municipality of Tibet (table 1). The Li et al study was conducted in Dangxiong County at an altitude of 4300 m, while the Cho et al study was conducted in Linzhou County (altitude 3650 m). The age of the subjects across the studies was between 19 and 82 years. The subjects included in the Cho et al study were normotensive and hypertensive, while the hypertensive status of the subjects included in the Li et al study was not specified.

Both studies measured the degree of agreement of automated BP devices (Omron HEM-759P and Omron HEM-7201) against the mercury sphygmomanometer. These devices had previously been validated for use at normal altitudes according to the validation requirements of the British Hypertension Society (BHS) and the Association for the Advancement of Medical Instrumentation (AAMI).

There were differences in the design and characteristics between the two studies (table 1). Li et al used a convenience sampling method to recruit patients’ relatives from a local community hospital and county government office employees (n=129). Cho et al recruited participants from a general outpatient clinic, but the sampling method was not specified (n=33). Neither study reported the socioeconomic or educational status of the participants. The Li et al study did not specify the protocol used for its validation, while Cho et al followed the European Society of Hypertension (ESH) International Protocol.

The procedures for BP measurements differed between the two studies (table 1). Li et al used three consecutive BP measurements in a seated position after 5 min rest with a 1 min interval between each measurement. Each measurement was taken simultaneously with mercury and electronic

Figure 1  A flow chart showing the process for identification and inclusion of studies examining the accuracy of blood pressure-measuring devices in Tibet.
sphygmomanometers connected by Y-tube by two trained observers (cardiologists) who were blinded to each other’s readings. Cho et al\textsuperscript{26} used 9–10 sequential measurements alternating between the mercury sphygmomanometer and automated device recorded by two trained observers (a physician and a medical student) who were blinded to each other’s readings. The interval allowed between the measurements was 30–60 s. Li et al\textsuperscript{25} used the mean of three procedurally sound BP measurements as the primary endpoint, but the method used to determine final BP measurement was not reported in Cho et al\textsuperscript{26}. The studies were funded by public institutions, and there were no reported conflicts of interests by the study authors.

There were some differences in the reporting quality of the included studies (figure 2A, B). Both studies did not specify whether they consecutively enrolled participants into their study (unclear risk of bias), and both studies avoided case–control designs (low risk). Li et al\textsuperscript{25} did not indicate whether there was an inappropriate exclusion in their study (unclear risk of bias), while Cho et al\textsuperscript{26} avoided that by including both hypertensive and normotensive patients in their study (low risk of bias).

Both studies\textsuperscript{25, 26} interpreted the result without considering the knowledge of the result of reference standard (low risk of bias). There was no prespecified threshold used in the Li et al study (high risk bias),\textsuperscript{25} but the Cho et al study used a prespecified threshold to test the performance.\textsuperscript{26} The reference standard used in both studies was likely to accurately classify the target condition (low risk of bias). Both studies interpreted the reference standards without knowledge of the result of the index test (low risk of bias). The flow and timing in both studies indicated low risks of bias when the interval between index test and reference standards was appropriate. All the patients received the same reference standards and the statistical analysis had included all the patients.

Regarding the concerns about the applicability, the participants included in both studies matched their patient selection criteria (low risk of bias; figure 2B); however, Li et al\textsuperscript{25} did not design the study based on any international validation standards for BP measurement, and the data therefore can be considered as insufficiently reported, which resulted in unclear applicability. Both studies showed low risk of bias in terms of the applicability for the reference standard (see e-appendix tables 1 and 2 for full rating of the quality of included studies).

Due to the limited number of included studies and the variation on methodology, a meta-analysis was considered inappropriate. Li et al\textsuperscript{25} reported a significantly higher combined SBP measurement with electronic sphygmomanometer compared with mercury sphygmomanometer (mean difference 5.8 ±4.7 mm Hg; p<0.001), with no significant difference in DBP measurement (mean difference 0.4±3.9 mm Hg; p=0.23; table 2). However, a strong linear relationship was shown between the measurements by the two devices with correlation coefficients of 0.97 and 0.96 for SBP and DBP, respectively. In the Cho et al study,\textsuperscript{26} the mean differences between the electronic and mercury sphygmomanometer measurements were 1.0 ±5.9 mm Hg for SBP and −3.1±4.6 mm Hg for DBP; p values or strengths of correlations between measurements were not reported. In total, 27 participants (82%) had measurements within 5 mm Hg in at least two out of three recordings for SBP with 25 participants (76%) for DBP; however, 3 participants had no measurements within 5 mm Hg for SBP and DBP.

In the Li et al study,\textsuperscript{25} the mean differences between calibrated readings between the mercury and electronic devices were −0.06 mm Hg (p=0.07) and 0.04 mm Hg (p=0.52) for SBP and DBP, respectively (table 2). In the Cho et al study,\textsuperscript{26} BP

<table>
<thead>
<tr>
<th>Study ID</th>
<th>Study design</th>
<th>Device tested</th>
<th>Location</th>
<th>Source of funding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Li et al\textsuperscript{25}</td>
<td>Observational Dangxiong County, Lhasa</td>
<td>Omron HEM-7201</td>
<td>Lhasa, Tibet, the US National Heart, Lung, and Blood Institute, and the UnitedHealth Chronic Disease Initiative</td>
<td>Bureau of Science and Technology of Tibet, the US National Heart, Lung, and Blood Institute, and the UnitedHealth Chronic Disease Initiative</td>
</tr>
<tr>
<td>Cho et al\textsuperscript{26}</td>
<td>Observational Linzhou County, Lhasa</td>
<td>Omron HEM-7201</td>
<td>Lhasa, Tibet, the US National Heart, Lung, and Blood Institute, and the UnitedHealth Chronic Disease Initiative</td>
<td>Bureau of Science and Technology of Tibet, the US National Heart, Lung, and Blood Institute, and the UnitedHealth Chronic Disease Initiative</td>
</tr>
</tbody>
</table>

**Table 1** Design and characteristics of studies examining the accuracy of automated BP monitors in Tibet

Due to the limited number of included studies and the variation on methodology, a meta-analysis was considered inappropriate.
readings were not calibrated; however, the authors noted that for 99 measurement pairs analysed, 72 and 68 device readings were within 5 mm Hg of observer measurements for SBP and DBP, respectively (90 and 92 for corresponding readings within 10 mm Hg).

In both studies, visual inspection of scatter plots of mean differences against mean BP measurements by both devices revealed that smaller proportions of participants with average SBP readings ≥140 mm Hg were enrolled into the studies (15%–16% for Li et al; actual values not reported for Cho et al). In the Li et al study, all the BP readings for this group of participants were within 1 SD of the mean SBP. In the Cho et al study, 94.6% of all BP recordings were within 2 SD of the mean difference between automated and mercury devices for SBP (95.3% for DBP).

In the Li et al study, the overall agreement rate was 85.3% and 93.8% for raw and calibrated readings of the electronic sphygmomanometers. The Cho et al study did not report the agreement rates between the electronic and mercury sphygmomanometers.

**DISCUSSION**

**Main findings**

The evidence from published studies suggests that electronic sphygmomanometers give similar BP measurements with the mercury sphygmomanometer when used to record DBP at high altitudes. The results of one study showed that electronic sphygmomanometer gave significantly higher SBP readings compared with mercury sphygmomanometer, while the findings of a second study showed similar measurements between both instruments. The results should be interpreted with caution because of the small number of studies included in the review and the variations in their designs and methodology. To our knowledge, this is the first systematic review, which assesses the degree of agreement of automated BP-measuring devices at high altitudes.

The 95% CIs for the mean difference between the automated and mercury sphygmomanometers observed in one study indicate that the BP measurements using the automated BP device are reliable, suggesting that the device can confidently be used to measure BP at high altitudes. Furthermore, the high agreement rates for SBP and DBP suggest that the differences observed in the mean BP readings between both devices were likely due to systematic rather than random errors. However, the lack of BP description of the participants creates some uncertainty about the trueness of the observed effect.

For over a century, the mercury sphygmomanometer was the gold standard for indirect measurement of BP. However, because of concerns about the environmental hazards associated with mercury, several other types of BP-measuring instruments have been developed over the past few years. Though many of them have been successfully integrated into clinical use in normal altitude settings, the results of our review suggest that there is limited evidence about their validity in Tibet, and possibly other settings at similar altitudes. The lack of validation

![Figure 2](https://example.com/image.png)

**Figure 2** (A): Risk of bias in included studies. (B) Applicability concerns of included studies. QUADAS, Quality Assessment of Diagnostic Accuracy Studies.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Main results of studies examining the accuracy of automated BP-measuring devices in Tibet*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study ID</strong></td>
<td><strong>Region</strong></td>
</tr>
<tr>
<td>Li et al25</td>
<td>Tibetan Autonomous</td>
</tr>
<tr>
<td>Cho et al26</td>
<td>Tibetan Autonomous</td>
</tr>
</tbody>
</table>

*SDs have been used as measures of dispersion where mean values are reported. BP, blood pressure; ES, electronic (automated) sphygmomanometer; MS, mercury sphygmomanometer.
studies creates some confusion as to whether the reportedly higher prevalence of hypertension in Tibet compared with mainland China (which is at a lower altitude) is accurate.

The density of mercury would not be expected to change at high altitudes because it is incompressible. It has been suggested that aneroid instruments are not affected by high altitude changes. However, the authors of one included study reported that the difference in systolic BP readings observed with the automated BP monitor could be due to changing altitudes, while those of the second study reported that the effect of altitude on automated BP monitors is unclear; therefore, this requires further investigation.

Comparison with existing literature
Our results are consistent with those of another study, which concluded that non-mercury sphygmomanometers (aneroid) might be suitable for use in measuring BP at high altitudes (Peru). In contrast to that report, the studies included in our review were confined to Tibetan inhabitants, and also tested the accuracy of electronic sphygmomanometers.

Strengths and limitations
This review has strengths. We used a robust method to search for eligible studies. We made efforts to contact the authors of the potential studies, and we accounted for the quality of included studies. In addition, both studies used trained and qualified personnel to perform BP measurements and recordings. However, we may not have identified all studies testing the accuracy of BP-measuring instruments in Tibet, especially unpublished studies. We included only two studies, and these differed in methodology. Because both studies used the Omron device, the study results cannot be generalised to other automated BP devices that use different algorithms. Furthermore, both studies were based on BHS grading, which is not designed for use at high altitudes. In addition, the included studies were conducted at altitudes below the average altitude in Tibet, and it is not known whether the results observed in these studies are applicable at altitudes above 4500 m; indeed, over half a million Tibetans live above 4500 m. These limitations prevent us from drawing firm conclusions about the degree of agreement of BP measurements with electronic sphygmomanometers at high altitudes.

Implications for research
More studies evaluating the validity of BP-measuring instruments at high altitudes, including settings over 4500 m above sea level, should be conducted. Future studies should also investigate the degree of agreement of such devices in children. Whether the European guideline requirements for sample size are appropriate for the Tibetan population is uncertain, and this warrants further investigation. Indeed, ESH guidelines suggest modification of sample size requirements for specific population groups; in addition, there are concerns that the current degree of agreement for automated sphygmomanometers specified in the guidelines is inadequate.

However, we note that the differences in BP recordings reported in both included studies fall within the limits observed in studies that compared automated devices with the mercury sphygmomanometer (eg, AAMI allows a difference of $\pm 8$ mm Hg for electronic or automated sphygmomanometers); consequently, it is possible that the readings at sea level for such comparisons may be similar.

CONCLUSION
The limited evidence from published studies suggests that automated (Omron) BP monitors show a high degree of agreement for DBP when compared against mercury sphygmomanometer at high altitudes. However, the degree of such agreement for SBP is not consistent. Few studies assessing the validity of automated BP monitors at high altitudes have been conducted, and they differ in design and methodology. All published studies until now have been conducted in adults. Further research assessing the suitability of BP-measuring instruments at high altitudes is therefore warranted.

**What is already known about this subject?**
At low altitudes, automated blood pressure (BP) monitor readings are highly correlated with standard mercury sphygmomanometer readings.

**What does this study add?**
At very high altitudes, automated (Omron) BP monitor readings are comparable with mercury sphygmomanometer readings for diastolic BP; however, systolic BP readings are not consistently correlated.

**How might this impact on clinical practice?**
BP measurement at high altitudes with automated devices may require calibration.

Acknowledgements
The authors thank Ms Nia Roberts for her help with conducting electronic searches.

Contributors
CM and IJO were involved with abstract screening, data extraction, data analysis and interpretation, and drafting of the review. CJH and AMW were involved with data analysis and interpretation, and drafting of the review.

Funding
CJH receives payment for running educational courses at the University of Oxford and University of Oxford ISIS consulting services for external teaching and training. He also receives royalties for books (Evidence Based Toolkit series by Blackwell BMJ Books).

Competing interests
None.

Provenance and peer review
Not commissioned; externally peer reviewed.

Open Access
This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

**REFERENCES**
Correction


The following should have been included in the Funding statement: CM is employed and supported through the ERC Starting Grant RATIMED (336932) at the Institute for Social Anthropology, Austrian Academy of Sciences.

*Heart Asia* 2016;8:76. doi:10.1136/heartasia-2016-010798corr1