What systematic reviews bring to the field of hypertension

Magali Leyvraz\textsuperscript{a}, Valérie Santschi\textsuperscript{b}, and Arnaud Chiolero\textsuperscript{a, c}

In the current issue of the \textit{Journal of Hypertension}, Duncombe \textit{et al.} [1] conducted a systematic review with meta-analysis to investigate the accuracy of oscillometric and aneroid blood pressure (BP) devices compared with the mercury sphygmomanometer for the measurement of BP in children. Following a well designed search strategy, and after a careful selection, they retained 38 studies including 26,879 children. Oscillometric devices yield slightly higher readings of SBP than auscultation with a mercury sphygmomanometer, with a difference of 1.6 mmHg (95\% confidence interval (CI): +0.6 to +2.8). The difference was smaller in the 12 studies using clinically validated devices (+1.8 mmHg; 95\% CI: +0.6 to +2.8). There was no statistically significant difference for DBP (+1.6 mmHg; 95\% CI: −0.2 to +3.3). However, there was a large heterogeneity between studies, explained only in part by differences in manufacturer, study setting, and observer training. Only three studies compared BP using aneroid and mercury device with, on average, comparable BP readings.

The current study is an excellent example of the growing number of systematic reviews in medicine and public health, particularly in the field of hypertension [2–5]. Initially, systematic reviews were designed to analyze together relatively small and inconclusive clinical trials [6,7]. Pooling data from these trials, to conduct a meta-analysis, results in a gain of statistical power, increasing the confidence in the effect estimates of the treatment. In cases of substantial clinical heterogeneity, however, pooling data and meta-analysis can be misleading [6,8] but, even without meta-analysis, systematic reviews are vital to have a comprehensive overview of all the available evidence on a given clinical question.

Nowadays, systematic reviews are not limited to the estimation of drug treatment effect using data from clinical trials. Indeed, they are useful, notably in the field of hypertension, to address a large variety of clinical questions, for example, to estimate the effect of complex clinical interventions to improve hypertension control [2,3], and using data from observational [9], measurement, or diagnostic studies [1,10]. Clinicians, health researchers, and policy makers are eager for high-quality systematic reviews. Although the number of systematic reviews and meta-analyses is increasing exponentially, a large proportion of these studies are of low quality and poor credibility [11,12]. Selected key features of a-priori high-quality reviews are listed in Table 1. Like for any clinical and epidemiological studies, a research protocol has to be written before the conduction of the review. Checklists of key items that the protocol should contain exist [13]. Public registration of the protocol, for example, in PROSPERO, is recommended to avoid duplicate reviews and, more importantly, to avoid selective reporting and prevent publication bias – a major plague of clinical and epidemiological research [13]. The identification of studies should be conducted through several electronic databases, using a systematic procedure, and other sources, such as trial registers and list of references of key articles to identify nonindexed studies [6,7]. The study selection and data extraction should be conducted in duplicate to limit errors and subjective decisions.

Once the data have been extracted, data synthesis and planned statistical analyses can be conducted. Assessment of the between studies heterogeneity is of major importance in systematic review; actually, describing and understanding the variability between studies may be the core of systematic reviews. If the studies included are sufficiently similar, data can be pooled for the meta-analysis. In the case of high heterogeneity, the reasons should be explored [14]. Quality assessment of included studies is essential to evaluate the confidence in the results of the meta-analysis. Meta-analyses including many low-quality studies will also be of low quality. Put more simply: 'Garbage in, garbage out'. Furthermore, publication bias has to be carefully evaluated. Indeed, because studies without statistically significant results or not going in the expected direction tend not to be published, the body of available evidence is biased [15]. Finally, the interpretation of the results should account for their level of credibility and confidence, and can be done using, for example, the Grading of Recommendations Assessment, Development, and Evaluation framework [16].
What systematic reviews bring to the field of hypertension

### TABLE 1. Key features of high-quality systematic reviews and meta-analyses, and their relations with the level of credibility and confidence in the results

<table>
<thead>
<tr>
<th>Key features</th>
<th>Increases credibility</th>
<th>Increases confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research protocol, published, or at least available on request</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Protocol written following, e.g., the Preferred Reporting Items for Systematic review and Meta-Analysis Protocols (PRISMA-P) guidelines [13]</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Explicit search strategy with several electronic databases and other sources</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Prospectively planned statistical analyses</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Registration of the protocol, e.g., on PROSPERO</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Study selection and data extraction in duplicate</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Evaluation of study quality</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Analysis of heterogeneity</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Assessment of publication bias</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Results reports following, e.g., the Preferred Reporting Items for Systematic review and Meta-Analysis (PRISMA) guidelines [18]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Duncombe et al. [1] through an explicit, well-documented, and exhaustive search identified a large number of studies comparing different BP measurement devices with auscultatory mercury sphygmomanometer among more than 25,000 children. Their major finding is that, on average, oscillometric devices yield slightly higher BP readings than by auscultatory method with a mercury sphygmomanometer. Significantly, when analyses were restricted to clinically validated oscillometric devices, the difference was smaller. Very few studies compared aneroid and mercury devices. Their review adds to the evidence that clinically validated oscillometric devices can be used to measure BP in children. It also demonstrates the paucity of data evaluating the aneroid method, despite their frequent use in pediatrics. Overall, the credibility of this review is high.

The high heterogeneity between studies the high probability of publication bias limit however our confidence in the estimates obtained by the meta-analysis. The authors found differences between measurement methods in SBP ranging from −5 to +15 mmHg. Differences in manufacturer or devices, study setting, and observer training explained only a part of this heterogeneity. The problem is that the larger the heterogeneity, the less confidence readers should have in the estimates obtained through meta-analysis [7]. It means that the pooled estimates and, more so, the CIs should be considered with caution. Rather than pooling data, the author could have reported only the average difference between measurement methods, with a measure of dispersion or variability (e.g., range), what would have been already highly informative.

A publication bias is a priori expected because studies of BP measurement devices with a large difference in readings may have not been published. The shape of the funnel plot is compatible with a publication bias, even if it could also reflect a true clinical heterogeneity between studies [17]. As a result, the differences between measurement methods (in one direction or the other) may have been underestimated. There is unfortunately no way to assess the exact impact of such bias.

Despite these limitations discussed in depth by Duncombe et al. [1], we would like to congratulate the authors for their high-quality systematic review. Their findings increase evidence-based knowledge in the field of BP measurement among children. We are further convinced, in agreement with the authors, that such studies help move hypertension guidelines from expert opinion toward evidence-based tools.

**ACKNOWLEDGEMENTS**

**Conflicts of interest**

There are no conflicts of interest.

**REFERENCES**

and apply the results to patient care: users’ guides to the medical literature. *JAMA* 2014; 312:171–179.


