Screening for Hypertension in Adults
Updated Evidence Report and Systematic Review for the US Preventive Services Task Force

Janelle M. Guirguis-Blake, MD; Corinne V. Evans, MPP; Elizabeth M. Webber, MS; Erin L. Coppola, MPH; Leslie A. Perdue, MPH; Meghan Soulsby Weyrich, MPH

IMPORTANCE Hypertension is a major risk factor for cardiovascular disease and can be modified through lifestyle and pharmacological interventions to reduce cardiovascular events and mortality.

OBJECTIVE To systematically review the benefits and harms of screening and confirmatory blood pressure measurements in adults, to inform the US Preventive Services Task Force.

DATA SOURCES MEDLINE, PubMed, Cochrane Collaboration Central Registry of Controlled Trials, and CINAHL; surveillance through March 26, 2021.

STUDY SELECTION Randomized clinical trials (RCTs) and nonrandomized controlled intervention studies for effectiveness of screening; accuracy studies for screening and confirmatory measurements (ambulatory blood pressure monitoring as the reference standard); RCTs and nonrandomized controlled intervention studies and observational studies for harms of screening and confirmation.

DATA EXTRACTION AND SYNTHESIS Independent critical appraisal and data abstraction; meta-analyses and qualitative syntheses.

MAIN OUTCOMES AND MEASURES Mortality; cardiovascular events; quality of life; sensitivity, specificity, positive and negative predictive values; harms of screening.

RESULTS A total of 52 studies (N = 215,534) were identified in this systematic review. One cluster RCT (n = 140,642) of a multicomponent intervention including hypertension screening reported fewer annual cardiovascular-related hospital admissions for cardiovascular disease in the intervention group compared with the control group (difference, 3.02 per 1000 people; rate ratio, 0.91 [95% CI, 0.86-0.97]). Meta-analysis of 15 studies (n = 11,309) of initial office-based blood pressure screening showed a pooled sensitivity of 0.54 (95% CI, 0.37-0.70) and specificity of 0.90 (95% CI, 0.84-0.95), with considerable clinical and statistical heterogeneity. Eighteen studies (n = 57,128) of various confirmatory blood pressure measurement modalities were heterogeneous. Meta-analysis of 8 office-based confirmation studies (n = 53,183) showed a pooled sensitivity of 0.80 (95% CI, 0.68-0.88) and specificity of 0.55 (95% CI, 0.42-0.66). Meta-analysis of 4 home-based confirmation studies (n = 1001) showed a pooled sensitivity of 0.84 (95% CI, 0.76-0.90) and a specificity of 0.60 (95% CI, 0.48-0.71). Thirteen studies (n = 5150) suggested that screening was associated with no decrement in quality of life or psychological distress; evidence on absenteeism was mixed. Ambulatory blood pressure measurement was associated with temporary sleep disturbance and bruising.

CONCLUSIONS AND RELEVANCE Screening using office-based blood pressure measurement had major accuracy limitations, including misdiagnosis; however, direct harms of measurement were minimal. Research is needed to determine optimal screening and confirmatory algorithms for clinical practice.
Hypertension is highly prevalent and one of the most important risk factors for cardiovascular disease (CVD).1-3 Blood pressure can be modified with lifestyle interventions,4-6 and good-quality randomized clinical trials (RCTs) demonstrate the effectiveness of antihypertensive pharmacological treatments to reduce CVD and total mortality.7,8 While office-based screening for hypertension in adults has been standard of care in the US for decades,9 office-based methods may misclassify individuals (white coat or masked hypertension). Contemporary research in blood pressure measurement has considered the potential benefits of out-of-office or novel office-based measurement modalities.

The aim of this updated systematic review was to inform an update of the 2015 US Preventive Services Task Force (USPSTF) recommendation on screening for hypertension in adults (A recommendation).10 This systematic review addressed the benefits and harms of screening for hypertension in adults, test accuracy of initial office-based screening measurements, and methods of confirmatory blood pressure measurement in those who initially screen positive.

Methods

Scope of Review

This review addressed 4 key questions (KQs) as shown in Figure 1. Methodological details including study selection, a list of excluded studies, additional data analysis methods, and sensitivity analyses are available in the full evidence report.11

Data Sources and Searches

MEDLINE, PubMed (publisher supplied records), the Cochrane Central Register of Controlled Trials, and CINAHL, were searched through August 17, 2019, to identify literature published after the previous review for the USPSTF12 (eMethods in the Supplement). The scope of this update differs from that of the 2015 review13 in that this review analyzed specificity and sensitivity of hypertension screening and confirmation, required ambulatory blood pressure measurement as the reference standard, included patients with diabetes, and did not address prognosis associated with various blood pressure measurement modalities. All included studies in the prior review and a subset of previously excluded studies were also evaluated, as well as reference lists of other systematic reviews and individual patient–data meta-analyses.13,14 ClinicalTrials.gov and the World Health Organization International Clinical Trials Registry Platform were searched for relevant ongoing trials. Active surveillance was conducted through March 26, 2021, via article alerts and targeted journal searches to identify major studies that might affect the conclusions or understanding of the evidence. No new studies were identified.

Study Selection

Investigators reviewed 21,741 unique citations and 544 full-text articles against a priori eligibility criteria (Figure 2 eTable 1 in the Supplement). All studies were required to enroll untreated adults or stratify results by treatment status and to have been conducted in countries rated as “very high” on the 2015 Human Development Index.15 Eligible populations for KQ2 (initial screening) were unselected based on blood pressure, whereas KQ3 populations (confirmatory screening) were preselected for having at least 1 elevated blood pressure measurement identified by clinic-based screening.

For KQ1 (screening), RCTs and nonrandomized controlled intervention studies were included that reported changes in health outcomes as a result of screening for hypertension compared with no screening. Eligible health outcomes were all-cause and cardiovascular mortality, cardiovascular disease events, symptomatic peripheral artery disease, vascular dementia, end-stage renal disease, and quality of life.

For KQs 2 and 3 (test accuracy), test accuracy studies comparing an initial office blood pressure measurement (OBPM) (KQ2) or confirmatory measurement modality (KQ3) with any ambulatory blood pressure monitoring (ABPM) reference standard were included. Attended and unattended automated office-based blood pressure (AOBP) measurements were eligible OBPM subtypes considered for all questions. The selection of the ABPM reference standard was based on a 2015 systematic review conducted for the USPSTF that concluded that ABPM was associated with cardiovascular events independently of OBPM and thus could serve as a reference standard.16 Other investigators have confirmed this finding.17

Confirmatory methods examined in KQ3 included repeated OBPM, self-OBPM (measurement performed by a patient in the office setting), home blood pressure measurement (HBPM), or kiosk. For KQ2a and KQ3a, included studies reported accuracy of protocol variations compared with an ABPM reference standard (eg, more vs fewer OBPM measures, more vs fewer days of HBPM). Studies needed to report sensitivity and specificity or provide enough data to calculate these values.

For KQ4 (harms), RCTs, nonrandomized controlled intervention studies, and cohort studies were included for the outcomes of quality of life, psychological effects of labeling, and absenteeism. Cross-sectional studies were additionally included for the outcome of ABPM tolerability.

Data Extraction and Quality Assessment

Two reviewers independently assessed the methodological quality of eligible studies. Disagreements were resolved by consensus and, if needed, consultation with a third reviewer. Each study was assigned a quality rating of “good,” “fair,” or “poor,” according to the USPSTF’s study design–specific criteria (eTable 2 in the Supplement).18 Studies rated poor quality because of serious methodological shortcomings were excluded.19 One reviewer abstracted descriptive and outcome data from each included study into standardized evidence tables; a second checked for accuracy and completeness.

Data Synthesis and Analysis

Results for KQ1 and KQ4 were analyzed qualitatively because of the small number of included studies reporting individual outcomes.

For test accuracy studies (KQ2 and KQ3), the primary outcomes of interest were sensitivity and specificity. For quantitative pooling, only studies that used both systolic blood pressure (SBP) and diastolic blood pressure (DBP) in their definition of hypertension were included because of relevance to current clinical practice. Because there is a lack of consensus on thresholds recommended by guidelines, thresholds were selected based on values most commonly reported in primary studies: 140/90 mm Hg for OBPM, 135/85 mm Hg for daytime ABPM, 130/80 for 24-hour ABPM, and 135/85 mm Hg for HBPM. Additional results for less commonly reported thresholds are available in the full evidence report.20 In quantitative analysis of KQ1 (initial screening), only studies measuring OBPM at a single visit were included; 2 additional studies measuring blood pressure at multiple visits were included in a sensitivity
Results for KQ3 (confirmatory measurement) were stratified by the type of confirmatory measure (repeat OBPM, HBPM, self-OBPM, AOBP, and kiosk). Data were sufficient for quantitative syntheses for OBPM and HBPM modalities only; other modalities were qualitatively synthesized. For all pooled analyses, a multivariate model was used to model sensitivity and specificity simultaneously, thus accounting for the correlation between these variables.

Stataversion15.1(StataCorp)wasusedforallanalyses.Allsignif-

cance testing was 2-sided, and results were considered statisti-
cally significant if the P value was .05 or less.

The aggregate strength of evidence was assessed for each KQ using the approach described in the Methods Guide for Effectiveness and Comparative Effectiveness Reviews, based on the number, quality, and size of studies and the consistency and precision of results between studies.21

Analysis.19,20 Results for KQ3 (confirmatory measurement) were stratified by the type of confirmatory measure (repeat OBPM, HBPM, self-OBPM, AOBP, and kiosk). Data were sufficient for quantitative syntheses for OBPM and HBPM modalities only; other modalities were qualitatively synthesized. For all pooled analyses, a multivariate model was used to model sensitivity and specificity simultaneously, thus accounting for the correlation between these variables.

Stata version 15.1 (StataCorp) was used for all analyses. All significance testing was 2-sided, and results were considered statistically significant if the P value was .05 or less.

The aggregate strength of evidence was assessed for each KQ using the approach described in the Methods Guide for Effectiveness and Comparative Effectiveness Reviews, based on the number, quality, and size of studies and the consistency and precision of results between studies.21

Benefits of Screening

Key Question 1. Does screening for hypertension in adults improve health outcomes?

There were no population-based trials comparing hypertension screening with no screening. One good-quality community-based cluster RCT (n = 140,642) conducted in Canada examined the effectiveness of a multicomponent CVD health promotion program on CVD health outcomes when hypertension screening was the primary intervention.47 The community clusters received either the Cardiovascular Health Awareness Program (CHAP) intervention or no intervention. In the CHAP communities, residents 65 years and older were invited to participate in community pharmacy-based blood pressure screenings using an automated instrument and complete a standardized risk profile. Participants received their risk profile, risk-specific educational materials, and local community resource information. At 1-year follow-up, the intervention communities had a reduction in the number of hospital admissions per 1000 for composite events (rate ratio, 0.91 [95% CI, 0.86-0.97]). There were 3.02 fewer annual hospital admissions for CVD per 1000 persons in the intervention group compared with the control group (intervention group, –2.25 per 1000 persons; control group, 0.77 per 1000 persons). There were no statistically significant differences in all-cause mortality among admitted residents (rate ratio, 0.98 [95% CI, 0.92-1.03]); intervention group, –1.47 per 1000 persons; control group, 0.77 per 1000 persons). There were no statistically significant differences in all-cause mortality among admitted residents (rate ratio, 0.98 [95% CI, 0.92-1.03]); intervention group, –1.47 per 1000 persons; control group, 0.77 per 1000 persons). There were no statistically significant differences in all-cause mortality among admitted residents (rate ratio, 0.98 [95% CI, 0.92-1.03]); intervention group, –1.47 per 1000 persons; control group, 0.77 per 1000 persons). There were no statistically significant differences in all-cause mortality among admitted residents (rate ratio, 0.98 [95% CI, 0.92-1.03]); intervention group, –1.47 per 1000 persons; control group, 0.77 per 1000 persons). There were no statistically significant differences in all-cause mortality among admitted residents (rate ratio, 0.98 [95% CI, 0.92-1.03]); intervention group, –1.47 per 1000 persons; control group, 0.77 per 1000 persons). There were no statistically significant differences in all-cause mortality among admitted residents (rate ratio, 0.98 [95% CI, 0.92-1.03]); intervention group, –1.47 per 1000 persons; control group, 0.77 per 1000 persons). There were no statistically significant differences in all-cause mortality among admitted residents (rate ratio, 0.98 [95% CI, 0.92-1.03]); intervention group, –1.47 per 1000 persons; control group, 0.77 per 1000 persons). There were no statistically significant differences in all-cause mortality among admitted residents (rate ratio, 0.98 [95% CI, 0.92-1.03]); intervention group, –1.47 per 1000 persons; control group, 0.77 per 1000 persons). There were no statistically significant differences in all-cause mortality among admitted residents (rate ratio, 0.98 [95% CI, 0.92-1.03]); intervention group, –1.47 per 1000 persons; control group, 0.77 per 1000 persons). There were no statistically significant differences in all-cause mortality among admitted residents (rate ratio, 0.98 [95% CI, 0.92-1.03]); intervention group, –1.47 per 1000 persons; control group, 0.77 per 1000 persons). There were no statistically significant differences in all-cause mortality among admitted residents (rate ratio, 0.98 [95% CI, 0.92-1.03]); intervention group, –1.47 per 1000 persons; control group, 0.77 per 1000 persons). There were no statistically significant differences in all-cause mortality among admitted residents (rate ratio, 0.98 [95% CI, 0.92-1.03]); intervention group, –1.47 per 1000 persons; control group, 0.77 per 1000 persons). There were no statistically significant differences in all-cause mortality among admitted residents (rate ratio, 0.98 [95% CI, 0.92-1.03]); intervention group, –1.47 per 1000 persons; control group, 0.77 per 1000 persons). There were no statistically significant differences in all-cause mortality among admitted residents (rate ratio, 0.98 [95% CI, 0.92-1.03]); intervention group, –1.47 per 1000 persons; control group, 0.77 per 1000 persons). There were no statistically significant differences in all-cause mortality among admitted residents (rate ratio, 0.98 [95% CI, 0.92-1.03]); intervention group, –1.47 per 1000 persons; control group, 0.77 per 1000 persons). There were no statistically significant differences in all-cause mortality among admitted residents (rate ratio, 0.98 [95% CI, 0.92-1.03]); intervention group, –1.47 per 1000 persons; control group, 0.77 per 1000 persons). There were no statistically significant differences in all-cause mortality among admitted residents (rate ratio, 0.98 [95% CI, 0.92-1.03]); intervention group, –1.47 per 1000 persons; control group, 0.77 per 1000 persons).
Several studies additionally reported accuracy for other thresholds,35,37,54,70,80 and 2 studies used SBP-alone or DBP-alone thresholds.36,78 Only 1 study reported accuracy for an OBPM threshold of 130/80 mm Hg or greater,36 the diagnostic threshold recommended in the 2017 American College of Cardiology/American Heart Association guideline.101 While all but 1 study78 reported that 24-hour ABPM was conducted, most (13) studies used daytime ABPM as a reference standard (eTable 4 in the Supplement). Only 1 study had low risk of bias for all domains and was rated as good quality.52,55 All other studies were rated fair quality and many had at least medium risk of bias for patient selection, conduct of the index test, and conduct of the reference test.

Meta-analysis of 15 studies using SBP/DBP thresholds and measuring blood pressure at a single visit (n = 13,091) showed a pooled sensitivity of 0.54 (95% CI, 0.37-0.70) and pooled specificity of 0.90 (95% CI, 0.84-0.95) (Figure 3; eTable 5 in the Supplement). Substantial clinical and methodologic heterogeneity among the included studies contributed to considerable statistical heterogeneity not explained by any single participant or test characteristic. Among this set of studies, positive predictive values and negative predictive values ranged widely, from 0.35 to 0.97 and from

Twenty fair- to good-quality studies (n = 12,614) examined the test accuracy of OBPM for initial screening for hypertension compared with ABPM (eTable 3 in the Supplement).10,19,20,23,26,28,32,35,38,42,43,45,46,49,54,56,61-63,67,68,70,75-78,80,86,87,95-97. Participants in the studies were primarily recruited from community-based samples. Only 5 were conducted in the US. Overall, participants represented a wide range of demographic and clinical characteristics, including a large range of blood pressures. The prevalence of hypertension as defined by ABPM in the included studies reflected population heterogeneity and ranged from 12.6%54 to 88.9%.70

Index test measurement protocols were heterogeneous and deviated somewhat from currently commonly performed protocols in US practice. Studies mostly used mercury sphygmomanometers with blood pressure measured by the manual auscultatory method, had participants rest for 5 minutes prior to measurement, and used the mean of multiple measurements (up to 9 measurements) at a single sitting (eTable 4 in the Supplement). Most other protocol characteristics were sparsely or variably reported. Studies most commonly used an office blood pressure of greater than 140/90 mm Hg or of 140/90 mm Hg or greater as the diagnostic threshold for the index test (17/20 studies).19,20,23,26,32,35,38,42,46,49,54,56,61,67,70,75,80,87,95,96
Figure 3. KQ2: Test Accuracy of Screening Office Blood Pressure Monitoring at a Threshold of ≥140/90 mm Hg to Identify Hypertension Detected by Ambulatory Blood Pressure Monitoring

<table>
<thead>
<tr>
<th>Source</th>
<th>Total No.</th>
<th>Hypertension, %</th>
<th>Positive screen, %</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abdalla et al, 23 2016</td>
<td>282</td>
<td>34.4</td>
<td>12.8</td>
<td>0.33 (0.24-0.43)</td>
<td>0.98 (0.95-0.99)</td>
</tr>
<tr>
<td>Fagard et al, 32 2005</td>
<td>243</td>
<td>31.3</td>
<td>42.8</td>
<td>0.71 (0.60-0.80)</td>
<td>0.70 (0.63-0.76)</td>
</tr>
<tr>
<td>Gourlay et al, 37 1993</td>
<td>66</td>
<td>28.8</td>
<td>30.3</td>
<td>0.74 (0.51-0.88)</td>
<td>0.87 (0.75-0.94)</td>
</tr>
<tr>
<td>Hansen et al, 38 2006</td>
<td>1385</td>
<td>46.7</td>
<td>31.6</td>
<td>0.56 (0.52-0.60)</td>
<td>0.90 (0.87-0.92)</td>
</tr>
<tr>
<td>Ishikawa et al, 46 2010</td>
<td>129</td>
<td>34.1</td>
<td>32.6</td>
<td>0.70 (0.56-0.82)</td>
<td>0.87 (0.78-0.93)</td>
</tr>
<tr>
<td>Kanno et al, 49 2010</td>
<td>775</td>
<td>13.9</td>
<td>19.4</td>
<td>0.48 (0.39-0.57)</td>
<td>0.85 (0.82-0.88)</td>
</tr>
<tr>
<td>Lyamina et al, 54 2012</td>
<td>269</td>
<td>12.6</td>
<td>0</td>
<td>0.00 (0.00-0.10)</td>
<td>1.00 (0.98-1.00)</td>
</tr>
<tr>
<td>Mancia et al, 55 2006</td>
<td>2024</td>
<td>33</td>
<td>42</td>
<td>0.74 (0.71-0.77)</td>
<td>0.74 (0.71-0.76)</td>
</tr>
<tr>
<td>O’Flynn et al, 57 2016</td>
<td>577</td>
<td>21</td>
<td>37.6</td>
<td>0.77 (0.69-0.83)</td>
<td>0.73 (0.69-0.77)</td>
</tr>
<tr>
<td>Poudel et al, 58 2019</td>
<td>432</td>
<td>88.9</td>
<td>62.5</td>
<td>0.68 (0.63-0.73)</td>
<td>0.83 (0.70-0.91)</td>
</tr>
<tr>
<td>Scuteri et al, 59 2016</td>
<td>2955</td>
<td>27.2</td>
<td>34.6</td>
<td>0.92 (0.90-0.94)</td>
<td>0.87 (0.86-0.88)</td>
</tr>
<tr>
<td>Shimbo et al, 60 2012</td>
<td>813</td>
<td>18.8</td>
<td>8.9</td>
<td>0.33 (0.26-0.40)</td>
<td>0.97 (0.95-0.98)</td>
</tr>
<tr>
<td>Thomas et al, 61 2017</td>
<td>441</td>
<td>31.7</td>
<td>14.3</td>
<td>0.31 (0.24-0.40)</td>
<td>0.94 (0.90-0.96)</td>
</tr>
<tr>
<td>Wei et al, 62 2016</td>
<td>717</td>
<td>15.5</td>
<td>10.5</td>
<td>0.34 (0.26-0.43)</td>
<td>0.94 (0.92-0.96)</td>
</tr>
<tr>
<td>Wojciechowska et al, 63 2016</td>
<td>201</td>
<td>32.8</td>
<td>30.3</td>
<td>0.62 (0.50-0.73)</td>
<td>0.85 (0.78-0.90)</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td></td>
<td></td>
<td>0.54 (0.37-0.70)</td>
<td>0.90 (0.84-0.95)</td>
</tr>
</tbody>
</table>

Sensitivity: 0.4 to 0.97, respectively. False-positive and false-negative rates likewise ranged widely (false-positive rate range, 0%-97%; false-negative rate range, 8%-100%). A sensitivity analysis adding 2 studies measuring blood pressure at multiple visits19,20 rendered the same point estimate but with slightly narrower confidence intervals (sensitivity, 0.53 [95% CI, 0.37-0.68]; specificity, 0.91 [95% CI, 0.85-0.95]).

Three additional studies (n = 1268) could not be included in the meta-analysis (eTable 5 in the Supplement). These included 1 study of attended AOBP25 with insufficient reporting for pooling showing sensitivity consistent with the pooled analysis but lower specificity (0.74 [95% CI, 0.66-0.82]) and 2 studies that used SBP-only or DBP-only thresholds.36,78

Four studies (n = 1467) reported results for multiple OBPM thresholds (eTable 5 in the Supplement).37,54,78,80 These studies consistently showed increased sensitivity and decreased specificity as thresholds are lowered. One study reported accuracy for an OBPM threshold of 130/80 mm Hg or greater in addition to 140/90 mm Hg or greater but also lowered the reference standard threshold; therefore, accuracy between the 2 OBPM thresholds cannot be directly compared.70 The resulting sensitivity for the OBPM threshold of 130/80 mm Hg or greater compared with the 130/80 mm Hg or greater daytime ABPM reference standard was 0.56 (95% CI, 0.50-0.61), with specificity of 0.89 (95% CI, 0.83-0.93).

Key Question 3a. What is the accuracy of confirmatory blood pressure measurement in adults who initially screen positive for hypertension compared with the reference standard (ABPM)?

Eighteen fair- to good-quality studies (n = 57 128) examined the diagnostic accuracy of confirmatory blood pressure measurements compared with an ABPM reference standard in adults with a previously detected elevated OBPM (eTable 6 in the Supplement).25,28,30,33,34,40,44,51,52,57,65,66,69,74,81,88,90,99 The Spanish ABPM Registry included 45 020 untreated individuals and represents much of the included evidence for this question.28 Only 2 studies were conducted in the US.30,44 Participants in the studies included patients referred by primary care physicians to blood pressure clinics because of borderline or elevated blood pressures, consecutive patients referred to ABPM or hypertension clinics, or individuals newly diagnosed as hypertensive by OBPM and not yet treated. Overall, the participants represented a wide range of demographic and clinical characteristics (eTable 6 in the Supplement). The prevalence of hypertension as defined by ABPM among this preselected population ranged from 47%74,99 to 80%.69 Two of the included studies were rated as good quality, with low risk of bias for all domains.65,90 All other studies were rated fair quality.

Four confirmatory blood pressure measurement modalities were examined for this KQ: repeated office blood pressure measurement (repeat OBPM), twice-daily home blood pressure measurement for 3 to 7 days (HBPM), measurement performed by a patient in the office setting (self-OBPM), and a truncated 6-hour ambulatory blood pressure measurement (truncated ABPM).

Key Question 3b. What is the accuracy of confirmatory blood pressure measurement modalities compared with each other and with ABPM?

Repeat OBPM

The majority of evidence (13/18 studies) was for repeat OBPM.23,33,34,40,44,51,52,57,65,81,88,90,99 As in KQ2, most OBPM confirmatory measurements were obtained with the patient seated with at least 5 minutes’ rest, attended by personnel, taken with a mercury
sphygmomanometer, used a diagnostic threshold of 140/90 mm Hg or greater, and were conducted at a single visit (eTable 7 in the Supplement). Other protocol details varied widely. Meta-analysis of 8 OBPM confirmation studies (n = 53,183) reporting SBP/DBP thresholds showed a pooled sensitivity of 0.80 (95% CI, 0.68-0.88) and a pooled specificity of 0.55 (95% CI, 0.42-0.66) with high heterogeneity (Figure 4; eTable 8 in the Supplement). Among these 8 studies, positive predictive values ranged from 0.61 to 0.88 and negative predictive values from 0.30 to 0.82. False-positive rates ranged from 15% to 65% and false-negative rates from 10% to 65%. Five studies did not contribute to the meta-analysis because they used SBP-only or DBP-only index thresholds, reference test thresholds, or both, that are not relevant to current clinical practice or did not provide sufficient data for pooling; these studies similarly reported large variations in accuracy (eTable 8 in the Supplement). 30,40,44,51,99 One study reported results for multiple OBPM thresholds with increased sensitivity and decreased specificity as thresholds were lowered. 34 No included study reported accuracy for an OBPM threshold of 130/80 mm Hg or greater.

HBPM
Four studies (n = 1,001) examined HBPM as a confirmatory method. 25,65,66,69 In these studies, participants were instructed to measure blood pressure for 3 to 7 days in the morning and evening in the seated position after a rest period of usually 5 minutes (eTable 9 in the Supplement). Meta-analysis of these 4 HBPM confirmation studies with a threshold of 135/85 mm Hg or greater (n = 1,001) showed a pooled sensitivity of 0.84 (95% CI, 0.76-0.90) and pooled specificity of 0.60 (95% CI, 0.48-0.71) (Figure 4; eTable 10 in the Supplement). Positive predictive values ranged from 0.68 to 0.94 and negative predictive values from 0.46 to 0.86. False-positive rates ranged from 22% to 50% and false-negative rates from 7% to 24%. Two studies reported accuracy for multiple HBPM thresholds. 25,69 These studies consistently showed increased sensitivity and decreased specificity as index test thresholds are lowered.

Self-OBPM
Two studies (n = 698) evaluated an index test in which a participant used an HBPM device to take their own blood pressure in an office setting (self-OBPM) (eTable 11 in the Supplement). While many fundamental device and protocol characteristics were similar among these studies, thresholds were not comparable, and measurements were unattended by staff in 1 study. Only 1 study used SBP/DBP thresholds relevant to current clinical practice and reported high sensitivity (0.92) and low specificity (0.25) (eTable 12 in the Supplement). The positive predictive value in that study was 0.59 and the negative predictive value was 0.72. The false-positive rate was 75% and the false-negative rate was 8%.

Truncated ABPM
One study (n = 263) reported the accuracy of a truncated (6-hour) ABPM compared with a full 24-hour ABPM test (eTable 13 in the Supplement). 30 Sensitivity and specificity were 0.94 and 0.76, respectively, for the subgroup (n = 126) for whom the ABPM indication was borderline hypertension (Figure 14 in the Supplement). Sensitivity and specificity were 0.89 and 0.70, respectively, for the subgroup (n = 137) with suspected white coat hypertension.

Comparative Accuracy
Two studies (n = 564) reported the accuracy of multiple confirmation methods against the same ABPM reference standard. 30,66 One study (n = 361) reported the accuracy of repeat OPBM and HBPM compared with a daytime ABPM reference standard. 66 Sensitivity was high and similar for both index tests (0.85 [95% CI, 0.80-0.88] for OBPM and 0.87 [95% CI, 0.83-0.91] for HBPM).
Table. Summary of Evidence

<table>
<thead>
<tr>
<th>Study design (No. of observations)</th>
<th>Summary of findings</th>
<th>Consistency and precision</th>
<th>Other limitations</th>
<th>Strength of evidence</th>
<th>Applicability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>KQ1: Screening</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Cluster RCT (0 new) (n = 140 642)</td>
<td>No trials examined the effectiveness of HTN screening alone vs no screening</td>
<td>Consistency NA, reasonably precise</td>
<td>Confounding from multicompartment intervention</td>
<td>Moderate for small benefit</td>
<td>Population: adults ≥65 y</td>
</tr>
<tr>
<td></td>
<td>One community-based cluster RCT of a multicompartment CVD health promotion trial reported a 9% reduction in the No. of CVD-related hospital admissions (rate ratio, 0.91 [95% CI, 0.86-0.97]) but no difference in all-cause mortality</td>
<td></td>
<td>Short 10-week intervention and 1-year follow-up duration</td>
<td></td>
<td>Intervention: community-based intervention (community pharmacy)</td>
</tr>
<tr>
<td><strong>KQ2: Diagnostic accuracy of initial OBPM</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20 Cross-sectional studies (20 new) (n = 12 614)</td>
<td>Meta-analysis of 15 studies using SBP/DBP thresholds and measuring blood pressure at 1 visit (n = 11 309) showed a pooled sensitivity of 0.54 (95% CI, 0.37-0.70) and a pooled specificity of 0.90 (95% CI, 0.84-0.95) with considerable heterogeneity</td>
<td>Inconsistent, imprecise</td>
<td>Heterogeneous group of studies in terms of population, measurement protocols, blood pressure thresholds</td>
<td>Low evidence for low sensitivity and adequate specificity</td>
<td>Population: general adult population</td>
</tr>
<tr>
<td></td>
<td>Repeat OBPM: Meta-analysis of 8 OBPM confirmation studies (n = 53 183) reporting SBP/DBP thresholds showed a pooled sensitivity of 0.80 (95% CI, 0.68-0.88) and a pooled specificity of 0.55 (95% CI, 0.42-0.66), with considerable heterogeneity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HBPM: Meta-analysis of 4 HBPM confirmation studies (n = 1001) showed a pooled sensitivity of 0.84 (95% CI, 0.76-0.90) and pooled specificity of 0.60 (95% CI, 0.48-0.71), with considerable heterogeneity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Self-OBPM: Meta-analysis of 4 OBPM confirmation studies (n = 1001) showed a pooled sensitivity of 0.84 (95% CI, 0.76-0.90) and pooled specificity of 0.60 (95% CI, 0.48-0.71), with considerable heterogeneity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ABCP: Meta-analysis of 4 ABCP confirmation studies (n = 1001) showed a pooled sensitivity of 0.84 (95% CI, 0.76-0.90) and pooled specificity of 0.60 (95% CI, 0.48-0.71), with considerable heterogeneity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>KQ2a: Diagnostic accuracy of different OBPM protocol characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Cross-sectional studies (4 new) (n = 1612)</td>
<td>Three studies addressed how number of measurements and visits influences accuracy and showed mixed results</td>
<td>Inconsistent, imprecise</td>
<td>Few studies overall; single studies evaluating different comparisons of comparative accuracy of number of visits and measurements, making conclusions difficult</td>
<td>Insufficient to evaluate any single protocol characteristic</td>
<td>Population: general adult population</td>
</tr>
<tr>
<td></td>
<td>Repeat OBPM: Four studies reported wide-ranging sensitivities (0.20-0.92) and specificities (0.25-0.97)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HBPM: Meta-analysis of 4 HBPM confirmation studies (n = 1001) showed a pooled sensitivity of 0.84 (95% CI, 0.76-0.90) and pooled specificity of 0.60 (95% CI, 0.48-0.71), with considerable heterogeneity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Self-OBPM: Meta-analysis of 4 OBPM confirmation studies (n = 1001) showed a pooled sensitivity of 0.84 (95% CI, 0.76-0.90) and pooled specificity of 0.60 (95% CI, 0.48-0.71), with considerable heterogeneity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>KQ3: Diagnostic accuracy of confirmatory screen</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 Cross-sectional studies (12 new) (n = 57 128)</td>
<td>Repeat OBPM: Meta-analysis of 8 OBPM confirmation studies (n = 53 183) reporting SBP/DBP thresholds showed a pooled sensitivity of 0.80 (95% CI, 0.68-0.88) and a pooled specificity of 0.55 (95% CI, 0.42-0.66), with considerable heterogeneity</td>
<td>Repeat OBPM: inconsistent and imprecise</td>
<td>Repeat office: heterogeneity in population recruitment, blood pressure measurement protocols, thresholds</td>
<td>Repeat OBPM: low for adequate sensitivity and low specificity</td>
<td>Population: adults referred for ABPM because of elevated office blood pressures or suspicious for white coat hypertension</td>
</tr>
<tr>
<td>Repeat OBPM: 13 studies (n = 55 759)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HBPM: 4 studies (n = 1001)</td>
<td>Repeat OBPM: Meta-analysis of 8 OBPM confirmation studies (n = 53 183) reporting SBP/DBP thresholds showed a pooled sensitivity of 0.80 (95% CI, 0.68-0.88) and a pooled specificity of 0.55 (95% CI, 0.42-0.66), with considerable heterogeneity</td>
<td>Repeat OBPM: inconsistent and imprecise</td>
<td>Repeat office: heterogeneity in population recruitment, blood pressure measurement protocols, thresholds</td>
<td>Repeat OBPM: low for adequate sensitivity and low specificity</td>
<td>Population: adults referred for ABPM because of elevated office blood pressures or suspicious for white coat hypertension</td>
</tr>
<tr>
<td>Self-OBPM: 2 studies (n = 698)</td>
<td>HBPM: Meta-analysis of 4 HBPM confirmation studies (n = 1001) showed a pooled sensitivity of 0.84 (95% CI, 0.76-0.90) and pooled specificity of 0.60 (95% CI, 0.48-0.71), with considerable heterogeneity</td>
<td>HBPM: inconsistent and imprecise</td>
<td>HBPM: low for adequate sensitivity and low specificity</td>
<td>HBPM: low for adequate sensitivity and low specificity</td>
<td>Population: adults referred for ABPM because of elevated office blood pressures or suspicious for white coat hypertension</td>
</tr>
<tr>
<td>Truncated vs 24-h ABPM: 1 study (n = 263)</td>
<td>Self-OBPM: Meta-analysis of 4 OBPM confirmation studies (n = 1001) showed a pooled sensitivity of 0.84 (95% CI, 0.76-0.90) and pooled specificity of 0.60 (95% CI, 0.48-0.71), with considerable heterogeneity</td>
<td>Self-OBPM: inconsistent and imprecise</td>
<td>Self-OBPM: insufficient</td>
<td>Self-OBPM: insufficient</td>
<td>Population: adults referred for ABPM because of elevated office blood pressures or suspicious for white coat hypertension</td>
</tr>
<tr>
<td>AOBP: no studies</td>
<td>Truncated ABPM: NA for consistency, precision</td>
<td>Truncated ABPM: NA for consistency, precision</td>
<td>Truncated ABPM: insufficient</td>
<td>Truncated ABPM: insufficient</td>
<td>Population: adults referred for ABPM because of elevated office blood pressures or suspicious for white coat hypertension</td>
</tr>
<tr>
<td><strong>(continued)</strong></td>
<td>Repeat office: heterogeneity in population recruitment, blood pressure measurement protocols, thresholds</td>
<td>Repeat office: heterogeneity in population recruitment, blood pressure measurement protocols, thresholds</td>
<td>Repeat office: heterogeneity in population recruitment, blood pressure measurement protocols, thresholds</td>
<td>Repeat office: heterogeneity in population recruitment, blood pressure measurement protocols, thresholds</td>
<td>Population: adults referred for ABPM because of elevated office blood pressures or suspicious for white coat hypertension</td>
</tr>
<tr>
<td></td>
<td>Repeat office: heterogeneity in population recruitment, blood pressure measurement protocols, thresholds</td>
<td>Repeat office: heterogeneity in population recruitment, blood pressure measurement protocols, thresholds</td>
<td>Repeat office: heterogeneity in population recruitment, blood pressure measurement protocols, thresholds</td>
<td>Repeat office: heterogeneity in population recruitment, blood pressure measurement protocols, thresholds</td>
<td>Population: adults referred for ABPM because of elevated office blood pressures or suspicious for white coat hypertension</td>
</tr>
<tr>
<td></td>
<td>Repeat office: heterogeneity in population recruitment, blood pressure measurement protocols, thresholds</td>
<td>Repeat office: heterogeneity in population recruitment, blood pressure measurement protocols, thresholds</td>
<td>Repeat office: heterogeneity in population recruitment, blood pressure measurement protocols, thresholds</td>
<td>Repeat office: heterogeneity in population recruitment, blood pressure measurement protocols, thresholds</td>
<td>Population: adults referred for ABPM because of elevated office blood pressures or suspicious for white coat hypertension</td>
</tr>
<tr>
<td>Key Question 3a. What confirmation protocol characteristics define the best test accuracy?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Five of 18 confirmation studies reported within-study comparisons of protocol characteristics on accuracy. Evidence on protocol variations for any one confirmation modality was sparse, but very limited evidence from 2 studies (n = 459) may suggest that for HBPM, additional days of measurement beyond 5 do not improve accuracy.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Harms of Screening

#### Key Question 4. What are the harms of screening for hypertension in adults?

Thirteen fair- to good-quality studies (n = 5150) examined the harms of screening and diagnosis of hypertension. Evidence for KQ4 is derived from heterogeneous populations and studies of limited quality largely performed 2 or more decades ago (eTable 15 in the Supplement). The limited existing evidence suggests that screening is associated with no decrement in quality of life or psychological distress and the scant evidence on screening's effect on absenteeism is mixed. ABPM follow-up testing is associated with minor adverse events including temporary sleep disturbance and bruising. Inaccurate diagnoses (false-positive and false-negative results) are also considered harms of screening and confirmation and have been discussed under KQ2 and KQ3 results.

### Discussion

This study reviewed the benefits and harms of screening for hypertension in adults, as well as the accuracy of tests; a summary of the evidence by key question is provided in the Table. The lack of contemporary population-based trials solely evaluating hypertension screening may be expected; such trials would not be considered feasible or ethical given that hypertension screening is standard practice and there is a robust evidence base linking asymptomatic hypertension treatment to improved CVD outcomes. Thus, the focus of this review was on the accuracy of screening (KQ2) and confirmatory (KQ3) blood pressure measurements, protocol variations that may influence accuracy (KQ2a and KQ3a), and the harms of screening and confirmation of hypertension (KQ4).

To our knowledge, this is the only published systematic review comparing the accuracy of office-based screening with an ABPM gold standard (KQ2). In the context of hypertension confirmation, the results of the present systematic review on the accuracy of confirmation (KQ3) are reasonably consistent with data from the International Database of Ambulatory Blood Pressure in relation to
Cardiovascular Outcome (n = 4997) and other systematic reviews of confirmation, even though other reviews have included mixed populations of treated and untreated individuals and populations with and without previous elevated OBPMs. The highly variable specificities in these reviews of confirmation likely reflect heterogeneity in populations and measurement protocols.

Any hypertension screening algorithm using measurement modalities other than ABPM alone will incur a considerable number of missed cases of masked hypertension as well as treatment of white coat hypertension. However, the clinical significance of the poor accuracy of OBPM is largely unknown. Subsequent consequences of poor OBPM accuracy could include delays in the identification and treatment of masked hypertension. For white coat hypertension, poor OBPM accuracy could result in unnecessary treatment and exposure to adverse effects or conversely a treatment benefit. Meta-analyses suggest that for untreated individuals generally recruited from population-based cohorts, cardiovascular risk progressively increases in the order of normotension, white coat hypertension, masked hypertension, and sustained hypertension. There are no clinical effectiveness trials for the treatment of masked hypertension, and subanalyses of trials analyzing the treatment benefit in white coat hypertension have yielded mixed results. Nonetheless, the robust evidence base supporting hypertension screening and treatment have historically been based solely on OBPM; therefore, participants with white coat hypertension were invariably included in those treatment trials.

Multiple strategies have been suggested to improve accuracy for identifying those with sustained and masked hypertension. AOBP has been suggested as a replacement for traditional office screening and out-of-office confirmation modalities. However, there were no included studies of unattended AOBP and only 1 study of attended AOBP reporting test accuracy compared with an ABPM reference standard. Other systematic reviews have suggested that, on average, mean AOBP and ABPM values in terms of mm Hg are similar; however, there is substantial heterogeneity and it is unclear if lack of mean mm Hg differences would result in similar diagnostic categorization and treatment decisions. Because higher 10-year CVD risk scores have been associated with an increased prevalence of masked hypertension, CVD risk tools could be useful for identifying specific populations that may benefit from ABPM to identify masked hypertension. Lowering the OBPM screening threshold is a possible approach to increase test sensitivity for sustained hypertension or to additionally identify a population for whom ABPM may be ordered to detect masked hypertension. Despite 2017 guidance from the American College of Cardiology/American Heart Association lowering the OBPM diagnostic threshold to 130/80 mm Hg or greater, no studies are available in an untreated population that report accuracy of this threshold compared with 140/90 mm Hg or greater using the same ABPM reference standard threshold. Trials examining the comparative accuracy and feasibility of various blood pressure measurement strategies for diagnostic confirmation of hypertension in primary care are needed; the publication of 1 such trial (BP-CHECK [NCT03130257]) is anticipated in 2021.

Limitations

This systematic review has several limitations. First, it excluded accuracy studies in which 20% or more of participants were treated to approximate screening populations. The accuracy of blood pressure measurements may be influenced by blood pressure variability, and variability may be reduced by hypertension medications. These pooled accuracy estimates therefore may not be applicable to treated populations. Second, for confirmatory test accuracy (KQ3), studies were included that enrolled participants referred for ABPM; while there are indications for ABPM referral outside of diagnostic confirmation, the lack of treatment was considered a proxy for diagnostic confirmation. Third, this review did not include accuracy studies that only reported mm Hg differences between measurement modalities or studies that only included κ values as a measure of agreement because clinical decision-making to initiate pharmacotherapy is based on blood pressures exceeding a defined threshold. Fourth, the reference standard for all accuracy studies was ABPM based on the previous review’s conclusion that there was a robust evidence base that ABPM is predictive of future CVD events; nonetheless, there is evidence suggesting that HBPM may be an alternative. Fifth, foundational evidence supporting screening is derived from treatment trials almost exclusively recruiting patients based on elevated office measurements without out-of-office confirmation. Sixth, treatment benefits and harms were beyond the scope of this review.

Conclusions

Screening using office-based blood pressure measurement had major accuracy limitations, including misdiagnosis; however, direct harms of measurement were minimal. Research is needed to determine optimal screening and confirmatory algorithms for clinical practice.

ARTICLE INFORMATION

Accepted for Publication: October 15, 2020.

Author Contributions: Dr Guirguis-Blake had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: Guirguis-Blake, Evans, Coppola, Weyrich.

Acquisition, analysis, or interpretation of data: Guirguis-Blake, Evans, Webber, Coppola, Perdue, Weyrich.

Drafting of the manuscript: Guirguis-Blake, Evans, Coppola.

Critical revision of the manuscript for important intellectual content: Guirguis-Blake, Evans, Webber, Coppola, Perdue, Weyrich.

Statistical analysis: Perdue.

Administrative, technical, or material support: Evans, Webber, Coppola, Weyrich.

Supervision: Guirguis-Blake, Evans.

Conflict of Interest Disclosures: None reported.

Funding/Support: This research was funded under contract HHSA-290-2015-000171-EPCS, Task Order 5, from the Agency for Healthcare Research and Quality (AHRQ), US Department of Health and Human Services, under a contract to support the US Preventive Services Task Force (USPSTF).

Role of the Funder/Sponsor: Investigators worked with USPSTF members and AHRQ staff to develop the scope, analytic framework, and key questions for this review. AHRQ had no role in study selection, quality assessment, or synthesis. AHRQ staff provided project oversight, reviewed the report to ensure that the analysis met methodological standards and distributed the draft for peer review. Otherwise, AHRQ had no role in the conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript findings.
on blood pressure, renin, aldosterone, catecholamines, cholesterol, and triglyceride. Cochrane Database Syst Rev. 2017;4:CD004022. doi:10.1002/14651858.CD004022.pub4


