Importance: Early identification of cognitive impairment may improve patient and caregiver health outcomes.

Objective: To systematically review the test accuracy of cognitive screening instruments and benefits and harms of interventions to treat cognitive impairment in older adults (≥65 years) to inform the US Preventive Services Task Force.

Data Sources: MEDLINE, PubMed, PsycINFO, and Cochrane Central Register of Controlled Trials through January 2019, with literature surveillance through November 22, 2019.

Study Selection: Fair- to good-quality English-language studies of cognitive impairment screening instruments, and pharmacologic and nonpharmacologic treatments aimed at persons with mild cognitive impairment (MCI), mild to moderate dementia, or their caregivers.

Data Extraction and Synthesis: Independent critical appraisal and data abstraction; random-effects meta-analyses and qualitative synthesis.

Main Outcomes and Measures: Sensitivity, specificity; patient, caregiver, and clinician decision-making; patient function, quality of life, and neuropsychiatric symptoms; caregiver burden and well-being.

Results: The review included 287 studies with more than 280 000 older adults. One randomized clinical trial (RCT) (n = 4005) examined the direct effect of screening for cognitive impairment on patient outcomes, including potential harms, finding no significant differences in health-related quality of life at 12 months (effect size, 0.009 [95% CI, −0.063 to 0.080]). Fifty-nine studies (n = 38 531) addressed the accuracy of 49 screening instruments to detect cognitive impairment. The Mini-Mental State Examination was the most-studied instrument, with a pooled sensitivity of 0.89 (95% CI, 0.85 to 0.92) and specificity of 0.89 (95% CI, 0.85 to 0.93) to detect dementia using a cutoff of 23 or less or 24 or less (15 studies, n = 12 796). Two hundred twenty-four RCTs and 3 observational studies including more than 240 000 patients or caregivers addressed the treatment of MCI or mild to moderate dementia. None of the treatment trials were linked with a screening program; in all cases, participants were persons with known cognitive impairment. Medications approved to treat Alzheimer disease (donepezil, galantamine, rivastigmine, and memantine) improved scores on the ADAS-Cog 11 by 1 to 2.5 points over 3 months to 3 years. Psychoeducation interventions for caregivers resulted in a small benefit for caregiver burden (standardized mean difference, −0.24 [95% CI, −0.36 to −0.13]) over 3 to 12 months. Intervention benefits were small and of uncertain clinical importance.

Conclusions and Relevance: Screening instruments can adequately detect cognitive impairment. There is no empirical evidence, however, that screening for cognitive impairment improves patient or caregiver outcomes or causes harm. It remains unclear whether interventions for patients or caregivers provide clinically important benefits for older adults with earlier detected cognitive impairment or their caregivers.
Dementia is a burdensome disease, not only to the health and longevity of individuals with the disease but also to their families and informal caregivers. According to the most recent Global Burden of Disease classification system, Alzheimer disease rose from the 12th most burdensome disease or injury in the United States in 1990 to the 6th in 2016 in terms of disability-adjusted life-years. It has been projected that by 2050 Alzheimer dementia will affect 13.8 million US residents. Early identification of cognitive impairment through screening would ideally allow patients and their families to receive care at an earlier stage in the disease process, potentially facilitating discussions regarding health, financial, and legal decision-making while the patient still retains decision-making capacity.

In 2014, the US Preventive Services Task Force (USPSTF) concluded that evidence was insufficient to assess the balance of benefits and harms of screening for cognitive impairment in older adults (I statement). The objective of this review was to inform an updated recommendation by the USPSTF.

Methods

Scope of Review
This review is an update of the 2013 review that supported the 2014 USPSTF recommendation. The update retained the analytic framework and key questions (KQs) that guided the 2013 review and included studies published since the previous review, as well as studies from the previous review that met updated inclusion criteria. No substantive changes were made to the scope of the review for this update, other than to exclude the medication tacrine from the list of included interventions as it is no longer available in the United States. The full report is available at https://www.uspreventiveservicestaskforce.org/Page/Document/UpdateSummaryFinal/cognitive-impairment-in-older-adults-screening. All main results presented in the full report are also presented in this article; more detailed methods and additional forest plots are included in the full report.

Data Sources and Searches
Ovid MEDLINE, PubMed (for publisher-supplied records only), PsycINFO, and the Cochrane Central Register of Controlled Trials were searched for relevant English-language literature (eMethods in the Supplement). Searches encompassed literature published through January 2019. The searches were supplemented by examining the reference lists of other previously published reviews and primary studies and by suggestions from experts. ClinicalTrials.gov was searched for ongoing randomized clinical trials (RCTs) related to KQ1. Active surveillance was conducted through November 22, 2019, 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
Because of the large volume of search results, we first used a single-screen process (ie, 1 reviewer screened for exclusion) for records with terms clearly outside the scope of the review in the title or abstract (eg, “mice,” “HIV,” “brain injury”). Two independent reviewers then screened the titles and abstracts and relevant full-text articles to ensure consistency with a prior inclusion and exclusion criteria (eTable 1 in the Supplement). For all KQs, studies that were relevant to community-dwelling, noninstitutionalized adults 65 years or older cared for in primary care in the United States were included. Only treatment studies (KQ4) conducted among community-dwelling older adults with mild cognitive impairment (MCI) or mild to moderate dementia were included (ie, those populations more representative of screen-detected older adults with cognitive impairment); studies of treatment of severe dementia were excluded.

For KQ1 and KQ3, we included RCTs and nonrandomized controlled studies that compared individuals who received screening with those who received no screening or usual care. For KQ2, studies that evaluated any brief screening instrument that could be administered by a clinician in 10 minutes or less or self-administered in 20 minutes or less were included. Studies of screening performed by diagnostic imaging or biomarker testing were excluded. Studies needed to report sensitivity and specificity (or data needed to calculate them) of a screening test compared with a diagnostic reference standard (ie, clinical assessment or neuropsychological testing with explicit diagnostic criteria with or without expert consensus/conference). Case-control studies and studies that selectively recruited patients with known or clinically suspected dementia or MCI (or control patients with normal cognition) were excluded because of the high risk of bias in patient selection for these studies.

For treatment effectiveness (KQ4), studies were limited to RCTs or nonrandomized controlled intervention studies of major pharmacologic and nonpharmacologic interventions intended for use during the early and mild stages of cognitive impairment and aimed at improving patient cognition, physical function, quality of life (QOL), caregiver burden or well-being, or a combination of these. Interventions with a primary aim of improving patient behavioral and psychological symptoms of dementia (eg, agitation, aggression, depressive symptoms), improving markers of physical performance, or reducing falls were excluded. Studies reporting outcomes on decision-making for patients, families, or clinicians (eg, health care planning, including advance directives; safety planning; legal and financial planning) were excluded. Studies required to report total adverse events, withdrawals attributable to adverse events, or serious adverse events that resulted in unexpected medical care, morbidity, or mortality.

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 study design–specific criteria (eTable 2 in the Supplement). Studies rated as of poor quality because of serious methodological shortcomings were excluded. One reviewer abstracted descriptive and outcome data from fair- and good-quality studies into standardized evidence tables; a second checked for accuracy and completeness.

Data Synthesis and Analysis
For test accuracy studies (KQ2), the primary outcomes of interest were sensitivity and specificity. Results were synthesized by instrument type (according to length of administration as very brief [administered in ≤5 minutes], brief [administered in 6-10 minutes], or longer [self-administered in >10 minutes] instruments) and separated by screening for dementia, MCI and dementia, or MCI only. Only 1 instrument had adequate data to conduct a quantitative synthesis: the Mini-Mental State Examination (MMSE) at a cutoff of 23 or less or 24 or less to detect dementia. A bivariate model was used to model sensitivity and specificity simultaneously, thus accounting for the correlation between these variables. For other instruments, ranges of sensitivity and specificity are reported.

For treatment studies, the interventions were grouped into 4 broad categories: (1) US Food and Drug Administration (FDA)–approved medications to treat Alzheimer disease (ie, acetylcholinesterase inhibitors [AChEIs] and memantine); (2) other medications or dietary supplements (eg, nonsteroidal anti-inflammatory drugs, gonadal steroids, and vitamins); (3) nonpharmacologic interventions aimed primarily at the patient, including cognitive training, stimulation, and/or rehabilitation, exercise interventions, and multicomponent and other interventions; and (4) nonpharmacologic interventions aimed primarily at the caregiver or caregiver-patient dyad, including psychoeducation, care and case management, and other caregiver-focused interventions.

Meta-analyses were conducted on the most commonly reported outcomes for each body of evidence. As a result, pooled analyses were conducted for FDA-approved medications on global cognitive function outcomes, global function outcomes, and harms; for nonpharmacologic patient-level interventions on global cognitive function outcomes; and for caregiver and caregiver-patient dyad interventions on caregiver burden and caregiver..
Depression measures. For consistency across the body of evidence, quantitative analyses focused on 6- to 12-month outcomes and included only shorter- or longer-term results when 6- or 12-month outcomes were not available.

Random-effects models using the DerSimonian and Laird method were used. For analyses with fewer than 10 studies, a sensitivity analysis was conducted using a more conservative restricted maximum likelihood analysis with the Knapp-Hartung correction. In cases in which continuous outcomes were measured using a variety of different instruments with differing scales (eg, caregiver burden), a standardized effect size (Hedges g) based on the differences in change between groups from baseline to follow-up was analyzed. A pooled risk ratio (for binary data) was used to analyze harms outcomes and improvement or maintenance in global function for AChEIs and memantine interventions.

The presence of statistical heterogeneity among the studies was assessed using standard χ² tests, and the magnitude of heterogeneity was estimated using the I² statistic. For outcomes with 10 or more studies in the meta-analysis, funnel plots were generated and an Egger test was conducted to evaluate small study effects and potential publication bias. Stata version 15.1 (StataCorp LP) 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.

Results

Investigators reviewed 11 645 unique citations and 967 full-text articles for all KQs (Figure 2). Overall, 287 studies including more than 285 000 older adults were included. Ninety-two studies were newly identified in this update and 195 were carried forward from the previous review. Fifty-nine studies that addressed the test accuracy of screening instruments (KQ2) were included, as well as another 224 RCTs and 3 observational studies that addressed the benefits and harms of screening or treatment (KQ1, KQ3, KQ4, and KQ5).

Benefits of Screening

Key Question 1. Does screening for cognitive impairment in community-dwelling older adults in primary care–relevant settings improve decision-making, patient-family/caregiver, or societal outcomes?

One RCT (IU CHOICE [conducted from October 2012 to September 2016]) examined the direct effect of screening for cognitive impairment on patient outcomes. This RCT was specifically designed and funded to address the lack of empirical data included in the previous USPSTF review. Primary care patients 65 years or older with no indication of cognitive impairment were randomized to screening for Alzheimer disease and related dementia (n = 2008) or no screening (n = 1997). Patients in the screening group were screened using the Memory Impairment Screen or the Mini-Cog and were referred for a voluntary diagnostic assessment if they screened positive on either or both tests. After a positive diagnostic assessment, a local memory care program worked with the caregivers and patients to provide or facilitate care and resources. Measures of health-related QOL using Health Utilities Index (HUI) scores (range, 0.36-1.00; 0 = dead and 1.00 = no impairment) were not significantly different between groups and across time. Among patients in the screening group, HUI scores were 0.67 (95% CI, 0.65 to 0.68) at baseline, 0.71 (95% CI, 0.69 to 0.72) at 1 month, 0.69 (95% CI, 0.66 to 0.71) at 6 months, and 0.68 (95% CI, 0.66 to 0.69) at 12 months. For those in the no screening group, HUI scores were not significantly different from scores in the screening group at all 4 time points (0.67 [95% CI, 0.66 to 0.69] at baseline, 0.69 [95% CI, 0.68 to 0.71] at 1 month, 0.70 [95% CI, 0.68 to 0.72] at 6 months, and 0.68 [95% CI, 0.66 to 0.70] at 12 months). Mixed-effects models showed no statistically significant differences between groups at any time point (eg, effect size at 12 months, 0.009 [95% CI, -0.063 to 0.080]). Furthermore, no significant differences in health care utilization, advance care planning, and dementia recognition by physicians were detected at 12 months.

Accuracy of Screening

Key Question 2. What is the accuracy of screening instruments to detect cognitive impairment in community-dwelling older adults?

Fifty-nine studies (n = 38 531) that addressed the test accuracy of screening for MCI or dementia were identified (eTable 3 in the Supplement). The number of participants screened ranged from 46 to 8805. Among the included studies, the prevalence of cognitive impairment varied widely: dementia ranged from 1% to 47%, MCI ranged from 10% to 52%, and cognitive impairment (inclusive of MCI and dementia) ranged from 17% to 90%.

The reference standard used to diagnose dementia or MCI usually consisted of a neuropsychological battery of tests and often was supplemented by a clinical examination, laboratory testing, imaging, assessment of depression and physical function, and/or an informant interview. The reference standard was administered by research staff, neurologists, psychiatrists, psychologists, psychometricians, other physicians, and/or nurses, and the diagnosis was usually made by consensus. Diagnostic and Statistical Manual of Mental Disorders (Fourth Edition, Third Edition Revised, and Third Edition) criteria were most often used to diagnose dementia, sometimes in conjunction with National Institute of Neurological and Communicative Disorders and Stroke–Alzheimer’s Disease and Related Disorders Association (NINCDS-ADRDA) criteria (for Alzheimer dementia) and National Institute for Neurological Disorders and Stroke–Association Internationale pour la Recherche et l’Enseignement en Neurosciences (NINDS-ARIE) criteria (for vascular dementia). No studies used Diagnostic and Statistical Manual of Mental Disorders (Fifth Edition) criteria. MCI was more variably diagnosed, with criteria including that from the International Working Group on MCI, performance 1 SD or more or 1.5 SD below normal, performance less than the 10th percentile on at least 1 cognitive test, a Clinical Dementia Rating scale score of 0.5, reported impairment that did not meet criteria for dementia, criteria developed by Petersen, criteria developed by a specific aging and disability resource center, or NINCDS-ADRDA criteria (for amnestic MCI).
The test accuracy of the MMSE to detect MCI was based on a cutpoint of 23 or less or 24 or less (eFigure 2 in the Supplement). The MMSE, a brief test that takes 7 to 10 minutes to complete, was the most-studied instrument (32 studies). The test performances of very brief and brief screening tests evaluated in only 1 study varied substantially (eFigures 4 and 5 in the Supplement). For self-administered, longer tests (>10 minutes), only 1 instrument (the Informant Questionnaire on Cognitive Decline in the Elderly) was assessed in more than 1 study, with sensitivity to detect dementia ranging from 0.74 to 1.0 and specificity ranging from 0.65 to 0.96 (eFigure 3 in the Supplement). For self-administered, longer tests (>10 minutes), only 1 instrument (the Informant Questionnaire on Cognitive Decline in the Elderly) was assessed in more than 1 study, with sensitivity to detect dementia ranging from 0.74 to 1.0 and specificity ranging from 0.65 to 0.96 (eFigure 3 in the Supplement). The test accuracy of 5 additional brief tests (7-Minute Screen, Abbreviated Mental Test, Montreal Cognitive Assessment, Saint Louis University Mental Status Examination, Functional Activities Questionnaire) was assessed in more than 1 study with sensitivity ranging from 0.74 to 0.88 and specificity ranging from 0.51 to 0.91 (eFigure 6 in the Supplement). Across all instruments, test performance was generally higher in the detection of dementia vs MCI, although confidence intervals overlapped.

Despite a large body of evidence examining cognitive screening instruments, most instruments were tested in only a few well-designed studies. The tests most likely relevant to screening in primary care are very brief instruments, with an administration time of 5 minutes or less. Eight very brief instruments were examined in more than 1 study (Clock Drawing Test, Lawton Instrumental Activities of Daily Living, Memory Impairment Screen, Mental State Questionnaire, Mini-Cog, verbal fluency tests, 8-item Interview to Differentiate Aging and Dementia [AD8], Functional Activities Questionnaire), with sensitivity to detect dementia usually at 0.75 or higher (range, 0.43-1.0) and specificity usually at 0.80 or higher (range, 0.54-1.0) (eFigure 1 in the Supplement). The MMSE, a brief test that takes 7 to 10 minutes to complete, was the most-studied instrument (32 studies). Pooled estimates across 15 studies (n = 12796) resulted in a sensitivity of 0.89 (95% CI, 0.85 to 0.92) and a specificity of 0.89 (95% CI, 0.85 to 0.93) of the MMSE to detect dementia at a cutpoint of 23 or less or 24 or less (eFigure 2 in the Supplement). The test accuracy of the MMSE to detect MCI was based on a much smaller body of literature (13 studies) with a variety of cutoffs and resulted in less consistent estimates for test accuracy, with a range in sensitivity from 0.20 to 0.93 and range in specificity from 0.48 to 0.93. The test accuracy of 5 additional brief tests (7-Minute Screen, Abbreviated Mental Test, Montreal Cognitive Assessment, Saint Louis University Mental Status Examination, Telephone Interview for Cognitive Status) was reported in more than 1 study, with sensitivity to detect dementia ranging from 0.74 to 1.0 and specificity ranging from 0.65 to 0.96 (eFigure 3 in the Supplement). The test performances of very brief and brief screening tests evaluated in only 1 study varied substantially (eFigures 4 and 5 in the Supplement). For self-administered, longer tests (>10 minutes), only 1 instrument (the Informant Questionnaire on Cognitive Decline in the Elderly) was assessed in more than 1 study, with sensitivity to detect dementia ranging from 0.80 to 0.88 and specificity ranging from 0.51 to 0.91 (eFigure 6 in the Supplement). Across all instruments, test performance was generally higher in the detection of dementia vs MCI, although confidence intervals overlapped.
Harms of Screening

Key Question 3. What are the harms of screening for cognitive impairment in community-dwelling older adults?

The IU CHOICE RCT (n = 4005) compared symptoms of depression and anxiety among patients randomized to screening for dementia vs those randomized to no screening.12,13 At 1 month after screening, depressive symptoms (as measured by the Patient Health Questionnaire 9 [PHQ-9]) and anxiety symptoms (as measured by the Generalized Anxiety Disorder 7 [GAD-7]) were not significantly different between groups after adjusting for baseline values. A similar pattern was evident at 6 and 12 months as well, suggesting no significant differences in feelings of depression or anxiety after screening for dementia.

Benefits of Interventions

Key Question 4. Do interventions for mild to moderate dementia or MCI in community-dwelling older adults improve decision-making, patient, family/caregiver, or societal outcomes?

Two hundred twenty-four RCTs77-300 representing more than 50,000 patients, caregivers, or both and 3 cohort studies301-303 with more than 190,000 patients were identified that addressed the treatment or management of MCI or mild to moderate dementia (Table 1; eTable 4 in the Supplement).

Acetylcholinesterase Inhibitors and Memantine

Based on 48 RCTs (n = 22,431) that evaluated AChEIs (ie, donepezil [18 RCTs; n = 6209], galantamine [10 RCTs; n = 7464], rivastigmine [8 RCTs; n = 4569]), and memantine (12 RCTs; n = 4189), these medications may improve measures of global cognitive function and global function in the short term (<6 months’ follow-up), but the magnitude of change was small (Table 1; eTable 5 in the Supplement).77,81,83,85,93,96,100,110,116,124,127,130,131,133-136

In meta-analyses, the differences in changes between patients receiving AChEIs or memantine compared with those receiving placebo ranged from approximately 1 to 2.5 points on the Alzheimer Disease Assessment Scale-Cognitive 11 (ADAS-Cog-11; scale range, 0-70) (n = 10,994) (eFigure 7 in the Supplement) and 0.5 to 1 point on the MMSE (scale range, 0-30) (n = 8589) (eFigure 8 in the Supplement) over 3 months to 3 years of follow-up. AChEIs and memantine appeared to increase the likelihood of improving or maintaining patients’ global function by 15% (for memantine) to 50% (for rivastigmine) over 3 to 12 months (pooled 95% CI range, 0.49 to 2.69) (n = 8405) (eFigure 9 in the Supplement); change at longer follow-up was not reported. Outcome measures of physical function were reported in only 60% of the studies and showed mixed results. Other important measures such as neuropsychiatric symptoms and rates of institutionalization were rarely reported; no medication studies included measures of QOL. Only 8 studies of medications examined outcomes beyond 6 months and generally found persistent effects that were consistent with shorter-term outcomes.

Most of the available evidence on the effectiveness of FDA-approved medications came from studies involving people with dementia, particularly among those with moderate vs mild forms of dementia, most commonly Alzheimer disease. Four RCTs (n = 1919; mean age, 74 years) tested these medications in people with MCI; these studies, testing donepezil and memantine, showed no benefit on global cognitive function. Only 1 RCT (n = 769) reported on progression of MCI to Alzheimer disease, finding no significant differences in the rate of conversion between people receiving donepezil vs placebo at 3 years.

Other Medications and Supplements

Twenty-nine RCTs (n = 6489; mean age, 75 years) evaluated other medications or supplements, including antihypertensives, 3-hydroxy-3-methyl-glutaryl-CoA reductase inhibitors (atorvastatin and simvastatin), nonsteroidal anti-inflammatory drugs (ibuprofen, naproxen, indomethacin, and celecoxib), gonadal steroids (estrogen [plus or minus progesterone] and testosterone), and dietary supplements and vitamins (multivitamins, B vitamins, vitamin E, and omega-3 fatty acids).78,79,114,120,121,122,123,128,155,156,182,195,210,213,223,226,227,234,247,248,254-256,258,260,264,277,285,300

There was no consistent benefit on global cognition or physical function in people with mild to moderate dementia or MCI for any of these medications or supplements (eTable 6 in the Supplement).

Nonpharmacologic Patient-Level Interventions

Sixty-one RCTs (n = 7847; mean age, 76 years) evaluated nonpharmacologic patient-level interventions, including cognitive-focused, exercise, and multicomponent and other interventions for people with MCI or dementia.80,82,84,88,90,91,94,98,101,102,104,106,117,119,123,133,152,158,161,164-166,169-171,174,175,177,180,185,187-189,191,197,212,215,217,219,220,222,229,232,233,235,237,244,251,253,257,263,265,270,272-274,278,280,282,297

In general, these studies were quite small and of limited duration.

Among all interventions, there was no clear benefit on global or domain-specific measures of cognitive function compared with control conditions at 3 months to 2 years of follow-up among people with MCI or dementia (Table 1; eTable 7 in the Supplement). Effect estimates generally favored intervention over control, but the magnitude of effect was inconsistent and had very wide confidence intervals (ranging from no effect to a large effect).

Although a pooled analysis of cognitive training, stimulation, and rehabilitation intervention studies found a small, statistically significant mean difference of 1.33 points on MMSE scores (95% CI, 0.29 to 2.37; 15 RCTs, n = 1341) favoring cognitive-focused interventions compared with control conditions at 3 to 12 months of follow-up, there was substantial clinical and statistical heterogeneity (eFigure 10 in the Supplement). Furthermore, combining 8 RCTs that reported changes in ADAS-Cog scores found a slightly greater improvement of 0.66 points (scale range, 0-70; higher scores indicate greater cognitive impairment) among intervention vs control group participants, but this difference was not statistically significant (mean difference, −0.66 [95% CI, −1.60 to 0.29]; n = 842) (eFigure 11 in the Supplement). There was no evidence that the effect of the interventions was modified by study, population, or intervention characteristics and no evidence of longer-term (up to 2 years) effects on cognitive function. Physical function outcomes, including change in activities of daily living and instrumental activities of daily living, as well as QOL and mental and neuropsychiatric symptoms, were inconsistently reported. Cognitive training, stimulation, and rehabilitation interventions consistently resulted in very little change over time or in small and relatively
equal decline in these measures from baseline to 3 months to 2 years across intervention and control groups, and few studies reported any statistically significant benefit. For RCTs of exercise interventions, pooled, conservative estimates of differences in measures of global cognitive function showed no to small effects based on the MMSE (mean difference, Table 1. Meta-analyses Results: Summary Across All Intervention Types (KQ4 and KQ5)
RCTs, n = 2776; standardized mean difference, –0.24 [95% CI, –0.36 to –0.13]; 27 significant benefit on caregiver burden at 3 to 12 months (standardized mean difference, –0.54 [95% CI, –0.96 to –0.12]; 8 RCTs, n = 1071) at 3 to 12 months (Table 1). There was, however, a pattern of effect for exercise interventions, with small improvements in measures of physical function and symptoms for intervention groups but declines for control groups. The clinical meaningfulness of these differences and the possibility of selective reporting limit the understanding of this finding. There was no consistent benefit of multicomponent and other patient-level interventions across outcomes.

Caregiver or Caregiver-Patient Dyad Interventions
Eighty-eight RCTs (n = 14 880; mean patient age, 78 years) evaluated the effect of multiple types of caregiver or caregiver-patient dyad interventions (Table 1). Overall, psychoeducation and care and case management interventions consistently benefited caregiver burden and depression outcomes (eTable 8 in the Supplement). Effect sizes were mostly small, however, and of unclear clinical significance. Psychoeducation interventions resulted in a small but statistically significant benefit on caregiver burden at 3 to 12 months (standardized mean difference, –0.24 [95% CI, –0.36 to –0.13]; 27 RCTs, n = 2776; I² = 50.2%) and in a medium effect on caregiver burden for care and/or case management interventions (standardized mean difference, –0.54 [95% CI, –0.96 to –0.12]; 8 RCTs, n = 1215; I² = 82.9%) (Table 1; eFigure 12 in the Supplement). The clinical importance of these changes in self-reported caregiver burden scores is unclear, with standardized effects translating to a between-group difference of approximately 2 to 4 points on the 22-item Zarit Burden Interview (Zarit-22; scale range, 0-88). Similar small effect sizes were seen for caregiver depression outcomes (eFigure 13 in the Supplement). The effect sizes of both caregiver depression and burden outcomes had wide confidence intervals, suggesting a range in the magnitude of benefit or, in some cases, a lack of benefit. There was no evidence in the meta-regression analyses that one type of intervention (psychoeducation vs care or case management vs other caregiver or caregiver-patient dyad interventions) was more effective than the others on measures of caregiver burden or caregiver depression. Likewise, there were no study, population, or intervention characteristics that consistently and robustly associated with larger effects on caregiver burden or depression outcomes.

Other outcomes such as caregiver or patient QOL, rates of or time to institutionalization, patient mental health and neuropsychiatric symptoms, and patient functional ability were sparsely reported across the studies, with no consistent evidence of benefit. Decision-making and preparation for meeting dementia-related needs were reported by only 1 RCT each (n = 414), with neither demonstrating statistically significant benefit.
### Table 2. Summary of Evidence

<table>
<thead>
<tr>
<th>Instrument or Treatment, Study Designs, Observations</th>
<th>Summary of Findings</th>
<th>Consistency and Precision&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Other Limitations</th>
<th>Strength of Evidence&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Applicability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>KQ1: Benefits of Screening</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 RCT (n = 4005)</td>
<td>No evidence of a difference in health-related QOL at 1, 6, and 12 mo between participants randomized to screening vs no screening as well as no significant difference in health care utilization and advanced planning at 12 mo</td>
<td>NA</td>
<td>High rate of missing data for all outcomes at all time points given attrition and data quality issues (42% missing data at 12 mo for primary outcome)</td>
<td>Low evidence of no benefit</td>
<td>Mean age of participants within the 1 RCT was 74.2 y, and the majority were women (66%) and white (67%) More than one-third of primary care-eligible older adults declined participation in the study, and 66% of those who screened positive refused further diagnostic assessment</td>
</tr>
<tr>
<td><strong>KQ2: Accuracy of Screening</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very brief instruments (31 cross-sectional studies [n = 22 359])</td>
<td>25 Instruments To detect dementia, sensitivity was usually at ≥0.75 and specificity at ≥0.80 Across all very brief instruments, the detection of MCI was less consistent, with a wide range in sensitivity and specificity</td>
<td>Reasonably consistent and precise (dementia); inconsistent and imprecise (MCI)</td>
<td>Large number of instruments with little replication</td>
<td>Moderate evidence of adequate sensitivity and specificity</td>
<td>Broad inclusion of older adult populations with a wide range of underlying dementia and MCI</td>
</tr>
<tr>
<td>Brief instruments (48 cross-sectional studies [n = 29 950])</td>
<td>20 Instruments For the MMSE, to detect dementia, 15 studies (n = 12 796) resulted in a pooled sensitivity of 0.89 (95% CI, 0.85 to 0.92) and a specificity of 0.89 (95% CI, 0.82 to 0.93) For other brief instruments reported in ≥1 study, sensitivity ranged from 0.74 to 1.0 and specificity ranged from 0.65 to 0.96 Across all brief instruments, the detection of MCI was less consistent, with a wide range in sensitivity and specificity</td>
<td>Reasonably consistent and precise (dementia); inconsistent and imprecise (MCI)</td>
<td>Large number of instruments with little replication, except for the MMSE</td>
<td>Moderate evidence of adequate sensitivity and specificity</td>
<td>Broad inclusion of older adult populations with a wide range of underlying dementia and MCI Administration time less useful for primary care screening</td>
</tr>
<tr>
<td>Longer, self-administered instruments (8 cross-sectional studies [n = 2271])</td>
<td>4 Instruments Only the IQCODE was assessed in ≥1 study, with sensitivity to detect dementia ranging from 0.80 to 0.88 and specificity ranging from 0.51 to 0.91 To detect MCI, sensitivity ranged from 0.71 to 0.82 and specificity ranged from 0.69 to 0.92</td>
<td>Reasonably consistent (dementia and MCI); precise (dementia and MCI)</td>
<td>Few instruments, little replication</td>
<td>Moderate evidence of adequate sensitivity and specificity</td>
<td>Broad inclusion of older adult populations with a wide range of underlying dementia and MCI</td>
</tr>
<tr>
<td><strong>KQ3: Harms of Screening</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continued)
Table 2. Summary of Evidence (continued)

<table>
<thead>
<tr>
<th>Instrument or Treatment, Study Designs, 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>1 RCT (n = 4005)</td>
<td>No evidence of a difference in symptoms of depression or anxiety between those in the screening vs no screening group at 1-, 6-, and 12-mo follow-up</td>
<td>NA</td>
<td>High rate of missing data for all outcomes at all time points given attrition and data quality issues (42% missing data at 12 mo for primary outcome)</td>
<td>Low evidence of no harm</td>
<td>Mean age of participants within the 1 RCT was 74.2 y, and the majority were women (66%) and white (67%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>KQ4: Benefits of Interventions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AChEIs and memantine (48 RCTs [n = 22 431])</strong></td>
</tr>
</tbody>
</table>

Pooled results indicate differences in change ranging from approximately 1 to 2.5 points in favor of drug groups on the ADAS-Cog-11 (range, 0-70)

Donepezil: MD, –2.13 (95% CI, –0.94 to –3.32); 6 studies (n = 1981); $I^2 = 64.4$

Galantamine: (MD, –2.13 (95% CI, –1.32 to –2.94); 9 studies (n = 3786); $I^2 = 65.9$

Rivastigmine: –2.43 (95% CI, –0.75 to –4.10); 5 studies (n = 2618); $I^2 = 81.9$

Memantine: –0.88 (95% CI, –0.11 to –1.65); 8 studies (n = 2609); $I^2 = 78.1$

Using accepted thresholds of clinical benefit, the average benefit across patients was not clinically significant

AChEIs and memantine increased the likelihood of improving or maintaining patients’ global function (eg, using a CIBIC+) by 15% (for memantine) to 50% (for rivastigmine) in the short term (pooled 95% CI range, 0.49 to 2.69)

Pooled change in global function found small effect sizes (SMDs ranging from 0.14 to 0.46)

Other important measures such as mental health and neuropsychiatric symptoms and rates of institutionalization were rarely reported; no RCTs included measures of QOL

**Other medications and supplements (29 RCTs [n = 6489])**

No evidence that antihypertensives, vitamins or omega-3 fatty acids, gonadal steroids, HMG-CoA reductase inhibitors, or NSAIDs are beneficial for any cognitive, functional, or other outcome at 3 mo to 4 y of follow-up | Reasonably consistent; imprecise | Small studies often with differential attrition between groups | Low evidence of no benefit | Older adults with mild to moderate dementia |

Unclear representation of ethnic minorities and those of varying education levels

(continued)
Table 2. Summary of Evidence (continued)

<table>
<thead>
<tr>
<th>Instrument or Treatment, Study Designs, 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>Nonpharmacologic patient-level interventions (61 RCTs [n = 7847])</td>
<td>No clear benefit of cognitive stimulation, training, or rehabilitation; exercise interventions; multicomponent interventions; and other interventions on global and domain-specific cognitive function compared with controls at 3 mo to 2 y follow-up among persons with MCI or dementia</td>
<td>Reasonably consistent; imprecise</td>
<td>Small studies of limited duration</td>
<td>Low evidence of small to no benefit</td>
<td>Broad range of older adults with MCI and mild and moderate dementia</td>
</tr>
<tr>
<td></td>
<td>Effect estimates generally favored intervention groups, but the magnitude of effects was inconsistent across RCTs and represented very wide confidence intervals</td>
<td></td>
<td>Types of outcomes, specific measures, and duration of follow-up was highly variable across studies</td>
<td></td>
<td>Very sparse reporting of clinical characteristics of the included patients such as race/ethnicity and education</td>
</tr>
<tr>
<td></td>
<td>Measures related to physical function, QOL, and mental and neuropsychiatric symptoms were only reported by one-half or less of the studies for each intervention group, and few found robust differences between groups</td>
<td></td>
<td></td>
<td></td>
<td>Virtually no data on effect modification by important clinical differences</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Many complex interventions may not be widely available in the United States</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caregiver and caregiver-patient dyad interventions (88 RCTs [n = 14 880])</td>
<td>Consistent benefit of psychoeducation and care and case management interventions on caregiver burden and depression outcomes; however, effect sizes were mostly small and are of unclear clinical significance</td>
<td>Reasonably consistent; precise</td>
<td>Little evidence of longer-term effects; inconsistency in outcomes and specific measures across studies, with many providing little data on precise scales used</td>
<td>Moderate evidence of small benefit</td>
<td>Generally applicable to caregivers of persons with moderate dementia</td>
</tr>
<tr>
<td></td>
<td>For caregiver burden, the standardized pooled effect was -0.24 (95% CI, -0.36 to -0.13); 27 studies (n = 2776); ( I^2 = 50.2% ) for psychoeducation interventions and -0.54 (95% CI, -0.83 to -0.22); 8 studies (n = 1215); ( I^2 = 82.9% ) for care and case management interventions</td>
<td></td>
<td></td>
<td></td>
<td>Many complex interventions may not be widely available in the United States</td>
</tr>
<tr>
<td></td>
<td>Other outcomes such as caregiver or patient QOL, rates or time to institutionalization, patient mental health and neuropsychiatric symptoms, and patient functional ability were sparsely reported across the RCTs with no consistent evidence of benefit</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Decision-making and preparation for meeting dementia-related needs were only reported by 1 RCT each, with neither finding statistically significant benefit of the interventions vs control conditions on overall scores for these measures</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

KQ5: Harms of Intervention

(continued)
Table 2. Summary of Evidence (continued)

<table>
<thead>
<tr>
<th>Instrument or Treatment, Study Designs, Observations</th>
<th>Summary of Findings</th>
<th>Consistency and Precision(^a)</th>
<th>Other Limitations</th>
<th>Strength of Evidence(^b)</th>
<th>Applicability</th>
</tr>
</thead>
<tbody>
<tr>
<td>AChEIs and memantine (48 RCTs ([n = 22,431]); 3 observational studies ([n = 190,076]))</td>
<td>Adverse effects from medications were common. Withdrawal or discontinuation was more common with AChEIs (13% withdrawing for donepezil and rivastigmine, 14% for galantamine) than placebo (8%). Memantine appeared to be better tolerated, with no significant difference in withdrawal rates (8%) compared with placebo (8%). In total, there did not appear to be a difference in total SAEs for these medications across studies with limited duration of follow-up; however, individual studies, including observational evidence, reported increased rates of bradycardia and, relatedly, of syncope, falls, and need for pacemaker placement among those exposed vs unexposed to AChEIs</td>
<td>Reasonably consistent; precise</td>
<td>The definitions of SAEs, which likely vary, were rarely described in the included studies</td>
<td>Moderate evidence of harm</td>
<td>Mostly represented patients with moderate dementia</td>
</tr>
<tr>
<td>Other medications and supplements (21 RCTs ([n = 5688]))</td>
<td>Across interventions, harms were not clearly significantly increased in intervention vs control groups</td>
<td>Reasonably consistent; precise</td>
<td>Small studies often with differential attrition between groups Lack of consistency in formulations and dosages of agents used</td>
<td>Low evidence of no harm</td>
<td>Older adults with mild to moderate dementia Unclear representation of ethnic minorities and those of varying education levels</td>
</tr>
<tr>
<td>Nonpharmacologic patient-level interventions (12 RCTs ([n = 2370]))</td>
<td>Little evidence of harms from good quality studies. Evidence of greater musculoskeletal problems among persons taking part in exercise interventions vs comparators</td>
<td>Reasonably consistent; precise</td>
<td>Sparse reporting of harms RCTs of exercise interventions more likely to report monitoring harms than cognitive training or other interventions</td>
<td>Low evidence of no harm(^b)</td>
<td>Applicable to patients with mild to moderate dementia and MCI</td>
</tr>
<tr>
<td>Caregiver and caregiver-patient dyad interventions (4 RCTs ([n = 486]))</td>
<td>No harms evident</td>
<td>NA</td>
<td>Sparse reporting of harms for patients or caregivers</td>
<td>Low evidence of no harm(^b)</td>
<td>Generally applicable to caregivers of persons with moderate dementia</td>
</tr>
</tbody>
</table>

Abbreviations: AChE, acetylcholinesterase inhibitor; ADAS-Cog, Alzheimer Disease Assessment Scale–Cognitive Subscale; CIBIC+, Clinician’s Interview-Based Impression of Change Plus Informant Input; HMG-CoA, 3-hydroxy-3-methyl-glutaryl-CoA reductase; IQCODE, Informant Questionnaire on Cognitive Decline in the Elderly; IQ, key question; MCI, mild cognitive impairment; MD, mean difference; MMSE, Mini-Mental State Examination; NA, not applicable; NR, not reported; NSAID, nonsteroidal anti-inflammatory drug; QOL, quality of life; RCT, randomized clinical trial; SAE, serious adverse event; SMD, standardized mean difference.

\(^a\) For KQ4, consistency, precision, and strength of evidence, assessments were based on primary outcomes within each body of evidence. For AChEIs and memantine, assessments were for cognitive function and global function outcomes. For other medications and supplements, assessments were for cognitive function. For nonpharmacologic patient-level interventions, assessments were for cognitive function, physical function, and neuropsychiatric symptoms. For caregiver and caregiver-patient dyad interventions, assessments were for caregiver burden and depression outcomes.

\(^b\) No hypothesized serious harms of nonpharmacologic patient or caregiver interventions. Thus, despite few studies reporting this outcome, there is low confidence that the finding of no harm in these RCTs reflects this body of evidence.
There is a large body of well-conducted test accuracy studies, but only a few instruments applicable to primary care have been examined in more than 1 study. Although the MMSE has the largest body of evidence to support its use and has adequate test accuracy, its utility is limited by the longer administration time (10-15 minutes) and cost (approximately $1.86 per form plus a test manual [$88], as of January 2020). Other instruments examined in at least 2 studies with adequate test performance to detect dementia among primary care-relevant populations include very brief instruments such as the Clock Drawing Test, the Memory Impairment Screen, the Mini-Cog, verbal fluency tests, the ADB, and the Functional Activities Questionnaire; brief instruments such as the Abbreviated Mental Test, Montreal Cognitive Assessment, 7-Minute Screen, and Saint Louis University Mental Status Examination; and the longer, self-administered Informant Questionnaire on Cognitive Decline in the Elderly.

One rationale for routine screening for cognitive impairment in older adults is facilitation of earlier diagnosis that may positively influence decision-making, leading to improved patient outcomes and reduced caregiver burden. This may include implementing medical, educational, and psychosocial interventions to suit individual patient and caregiver needs and encouraging patient participation in medical, legal, and financial decisions. While these are logical arguments, there is currently little empirical evidence, including qualitative evidence, to support them.

Screening for cognitive impairment may have direct or indirect harms as a result of diagnostic inaccuracy (false-positive and false-negative results) or negative emotions and stigma that may arise with diagnosis. Recent systematic reviews regarding patients’ attitudes and preferences about screening for dementia found mixed evidence. Some studies suggested that patients have no concerns, whereas others suggested that few people would agree to routine screening for memory problems for reasons such as stigma. Evidence suggests that caregivers and the general public believe they will benefit from being screened for dementia, in part because they believe there are effective treatments and financial benefits.

This review was not a comprehensive synthesis of all treatment and management options for people with cognitive impairment; instead, the focus was on selected interventions aimed at people with mild to moderate dementia or MCI. Based on the large body of evidence, there is support that AChEIs (donepezil, galantamine, and rivastigmine) and memantine and interventions that support caregivers, including care coordination, can result in small improvements in patient and caregiver health outcomes in the short term. The average effects of these benefits are quite small and likely not clinically significant. Any benefits are further limited by the commonly experienced adverse effects of medications and the limited availability of complex caregiver and care coordination interventions. Cognitive stimulation and training, exercise interventions, and other medications and supplements showed some favorable effects on patients’ cognitive and physical function, but study evidence lacked consistency and the estimates of benefit were imprecise.

Limitations

There is a lack of evidence around how screening for and treating MCI and early-stage dementia affects decision-making outcomes. Furthermore, there has been little reproducibility in testing specific screening instruments in primary care populations. The treatment literature is limited by a lack of consistency in the specific outcomes reported and short follow-up duration. It is difficult to interpret the clinical importance of the small average effects seen among treatment studies, and many measures likely have limited responsiveness for patients with less pronounced cognitive impairment. Consistent and standardized reporting of results according to meaningful thresholds of clinical significance would be helpful in interpreting the small average effects on continuous outcome measures. Other important measures such as QOL, physical function, and institutionalization were inconsistently reported.

Conclusions

Screening instruments can adequately detect cognitive impairment. There is no empirical evidence, however, that screening for cognitive impairment improves patient or caregiver outcomes or causes harm. It remains unclear whether interventions for patients or caregivers provide clinically important benefits for older adults with earlier detected cognitive impairment or their caregivers.
Institute on Aging and National Institute of Mental Health). Comments from reviewers were presented to the USPSTF during its deliberation of the evidence and were considered in preparing the final evidence review.

Editorial Disclaimer: This evidence report is presented as a document in support of the accompanying USPSTF Recommendation Statement. It did not undergo additional peer review after submission to JAMA.

REFERENCES


27. Ayoun L. The IQCODE versus a single-item informant measure to discriminate between cognitively intact individuals and individuals with dementia.
Clinical Review & Education
US Preventive Services Task Force

USPSTF Report: Screening for Cognitive Impairment in Older Adults


139. Gittin LN, Winter L, Burke J, Chernett N, Dennis MP, Hauck WW. Tailored activities to manage neuropsychiatric behavior in persons with dementia and reduce caregiver burden.


randomised controlled trial. 


194. Meeuwsen EJ, Melis RJ, Van Der Aa GC, et al. Maintenance cognitive stimulation therapy for memory clinicians or general practitioners: randomised controlled trial. BJM. 2013;347:f6276. doi:10.1136/bmj.f6276


211. Ostwald SK, Hepburn KW, Caron W, Burns T, Mantell R. Reducing caregiver burden.


Efficacy of multivitamin supplementation treatment in people with cognitive impairment. 195-203. doi:10.1002/gps.4312

Coaching intervention for neurocognitive disorder 521-531. doi:10.1097/JGP.0b013e3182110599


308. Le Couteur DG, Doust J, Creeze H, Brayne C. Political drive to screen for pre-dementia: not evidence based and ignores the harms of diagnosis. BMJ. 2013;347:f5125. doi:10.1136/bmj.f5125

