APPENDIX A. SEARCH STRATEGY

Symbol	Concept	Search Strategy (PubMed)
D	Dementia	"dementia"[MeSH Terms] OR "dementia"[All Fields]
SS	Signs + Symptoms	(((((("signs + symptoms"[MeSH Terms] OR ("signs"[All Fields] + "symptoms"[All Fields]) OR "signs + symptoms"[All Fields]) OR (warning sign[All Fields] OR warning signal[All Fields] OR warning signal/precue[All Fields] OR warning signals[All Fields] OR warning signs[All Fields] OR warning signs/symptoms[All Fields])) OR (red flag[All Fields] OR red flagging[All Fields] OR red flags[All Fields])) OR presenting[All Fields]) OR (suspect[All Fields] OR suspected[All Fields])) OR (predict[All Fields] OR predictor[All Fields] OR predictors[All Fields])
CS	Cross-Sectional Studies	"Cross-Sectional Studies"[Mesh]
SR	Systematic Review Subset	systematic[sb]
G	Guidelines, Consensus Statement Publication Type	Guideline [pt] OR Consensus Development Conference [pt]
T	Specific Tests	(("montreal cognitive assessment"[tiab])) OR (("moca"[tiab])) OR (("slums"[tiab])) OR (("st louis university mental status"[tiab])) OR (("saint louis university mental status"[tiab])) OR (("short test of mental status"[tiab])) OR (("STMS"[tiab])) OR (("General Practitioner Assessment of Cognition"[tiab])) OR (("GPCog"[tiab])) OR (("mini-cog"[tiab])) OR (("mini cog"[tiab])) OR (("orientation memory concentration"[tiab])) OR (("bomc"[tiab]))) NOT ((poverty areas[mesh]))

All searches were performed in July of 2009

Dementia Review #1 Key Question #1
PubMed
Primary Studies
(D + SS + CS) = 518
Secondary Studies (systematic reviews, guidelines or consensus statements only not general reviews)
((D + SS) + (SR OR G)) = 322

Additional databases

Cochrane central register of controlled trials and database of abstracts of reviews of effects

1 dementia.mp.
2 (signs and symptoms).mp
3 warning sign.mp.
4 warning signal.mp.
5 red flag*.mp.
6 presenting.mp.
7 suspect.mp.
8 suspected.mp.
9 predict*.mp
10 8 or 6 or 4 or 3 or 7 or 9 or 2 or 5
11 1 and 10

229 results after de-duplication 201

CINAHL

S1 ("dementia") or (MH "Dementia+")
S2 ("signs and symptoms") or (MH "Signs and Symptoms (Non-Cinahl)")
S3 "warning signs"
S4 "warning signal"
S5 "red flag*"
S6 "presenting"
S7 "suspect"
S8 "suspected"
S9 "predict"
S10 predictor*
S11 S2 or S3 or S4 or S5 or S6 or S7 or S8 or S9 or S10
S12 ("dementia") or (MM "Dementia+")
S13 S11 and S12
S14 S11 and S12 Narrow by Subject: Major Heading0: - Dementia

367 results after de-duplication 309

PsychINFO

1 exp *Dementia/ 30642
2 (signs and symptoms).mp. [mp=title, abstract, heading word, table of contents, key concepts] 5567
3 warning sign.mp. or exp Warnings/ 738

4 red flag.mp. 42
5 red flags.mp. 77
6 exp Symptoms/ or presenting.mp. 131298
7 suspect.mp. 1854
8 suspected.mp. 4720
9 predict.mp. 36261
10 predictor.mp. 27521
11 predictors.mp. 38646
12 6 or 11 or 3 or 7 or 9 or 2 or 8 or 4 or 10 or 5 222250
13 1 and 12 4251
14 limit 13 to ("0800 literature review" or "0830 systematic review" or 1200 meta analysis)

208 Results after de-duplication 192

AGELINE

Title: dementia
AND
Title: "signs and symptoms" ; "warning sign" ; "warning signal" ; "red flag" ; "red flags" ; presenting ; suspect ; suspected ; predict ; predictor

39 items after de-duplication 30 unique

We also re-executed the search described in Mitchell, 2008 in Medline.

(subjective memory OR memory complaint* OR memory difficult* [abstract]) AND (Dementia OR Alzheimer* OR mild cognitive [abstract]) AND (validity OR diagnosis OR sensitivity OR specificity OR accuracy OR re receiveOperator OR ROC [full text]) limited to 2008-Sept 2009 (date of search)

79 Results

Dementia Review #1 Key Question #2 & 3

PubMed
(D + T) = 54

Additional databases (named tests were searched in the following databases)
Cochrane central register of controlled trials and database of abstracts of reviews of effects: 27 Results, after de-duplication 0
HAPI: 24 Results, after de-duplication 23
PsycINFO: 87 Results, after de-duplication 74
CINAHL: 48 Results, after de-duplication 44

APPENDIX B. INCLUSION/EXCLUSION CRITERIA

Author, Year_____ Title or ID#_____

1. Do any of the following constitute the study population:
   a. A community sample of persons age 60+ that includes non-demented individuals as well as patients with mild to moderate dementia ..... ❑
   b. Patients with newly diagnosed, mild to moderate dementia ..... ❑
   c. Non-demented individuals aged 60+..... ❑
   d. None of the above.....STOP

2. Does the study use a standard criterion for diagnosing dementia?
   a. No .....STOP
   b. Yes (e.g. DSM-IV)..... ❑

3. Does the study evaluate any of the following cognitive tests:
   Blessed Orientation-Memory-Concentration Test (BOMC) ..... ❑
   Mini-Cog ..... ❑
   General Practitioner Assessment of Cognition (GPCOG)..... ❑
   Short Test of Mental Status (STMS) ..... ❑
   St. Louis University Mental Status Exam (SLUMS) ..... ❑
   Montreal Cognitive Assessment (MoCA) ..... ❑
   None of the above.....proceed to Q4

4. Does the study provide prevalence or other descriptive data on signs and/or symptoms of dementia?*
   a. No .....STOP
   b. Yes ..... ❑

5. Is the text of the article in English?
   a. No .....STOP
   b. Yes ..... ❑

6. If this article meets no other criterion, should it be saved for background or discussion?
   a. No .....STOP
   b. Yes: clinical guidelines..... ❑
   c. Yes: narrative review with potentially useful references..... ❑
   d. Yes: qualitative study discussing relevant signs and symptoms..... ❑
   e. Yes: other, specify..... ❑

Key words or categories:

Notes

Circle the Key Question(s) to which this article applies:

1. What signs and symptoms should prompt VA providers to assess cognitive function as part of an initial diagnostic workup for dementia?
2. Which measures of cognitive function provide the optimal sensitivity, specificity, and time to administer, and are readily available within the VA?
3. What are adverse consequences of using these measures?

** Examples of data that do not meet criteria for item 4 include epidemiologic risk factors for dementia; diagnostic imaging, laboratory, or physiological tests (e.g. sense of smell; cerebrospinal fluid studies; indicators of acute confusion; predictors of DNR orders*

APPENDIX C. QUADAS CRITERIA FOR EVALUATING DIAGNOSTIC ACCURACY STUDIES

Table 2: The Quadas tool

Item		Yes	No	Unclear
1.	Was the spectrum of patients representative of the patients who will receive the test in practice?	( )	( )	( )
2.	Were selection criteria clearly described?	( )	( )	( )
3.	Is the reference standard likely to correctly classify the target condition?	( )	( )	( )
4.	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	( )	( )	( )
5.	Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?	( )	( )	( )
6.	Did patients receive the same reference standard regardless of the index test result?	( )	( )	( )
7.	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	( )	( )	( )
8.	Was the execution of the index test described in sufficient detail to permit replication of the test?	( )	( )	( )
9.	Was the execution of the reference standard described in sufficient detail to permit its replication?	( )	( )	( )
10.	Were the index test results interpreted without knowledge of the results of the reference standard?	( )	( )	( )
11.	Were the reference standard results interpreted without knowledge of the results of the index test?	( )	( )	( )
12.	Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	( )	( )	( )
13.	Were uninterpretable/ intermediate test results reported?	( )	( )	( )
14.	Were withdrawals from the study explained?	( )	( )	( )

Source: Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J, Whiting P, et al. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Medical Research Methodology 2003;3:25. Used with permission.

APPENDIX D. PEER REVIEW COMMENTS

Reviewer	Comment	Response
Question 1. Are the objectives, scope, and methods for this review clearly described?		
1	Yes. Objective scope and methods are clearly described. The report is very succinct and to the point. However, I did not see QUADAS criteria fully explained. This should be explicitly described so a reader can independently evaluate how you did your evaluation.	Noted. We have clarified in the Methods that the details of the QUADAS criteria are listed in Appendix C.
2	Yes. Would be very helpful to have a glossary of terms to help non-expert readers (e.g., many policy-makers) understand the statistical terms (e.g., sensitivity, specificity, prevalence, incidence, positive/negative predictive value, screening, etc.) and other technical terms (e.g., stereognosis and graphesthesia). It is important that the methods, results, and discussion/conclusions/recommendations be clear in layperson as well as technical terms.	Will include glossary and we've tried to modify the language.
3	Yes. (No comment)	Noted.
4	Yes. Did you consider examining warning signs in combination rather than singly (e.g., memory complaint in combination with behavioral symptom such as apathy or driving violation). This may improve the discriminative properties of the warning signs. If you have not considered this – would you consider examining these warning signs in combination? That is a likely scenario for clinical use. It is unclear to me the extent to which you considered studies that discussed informants identification of (sic)	We've actually added some studies and narrative re: informant report. We added the suggestion re: consideration of examining these warning signs in combination to our future studies section.
5	Yes. I believe that this review sheds light on important questions and is very clear in the way it answers these questions with evidence based response/discussion. I found it to be very informative in looking at a large body of literature to answer specific questions that are very clinically relevant in dementia diagnosis/ recognition.	Noted.
Question 2. Is there any indication of bias in our synthesis of the evidence?		
1	There is no evident bias in synthesis of evidence however, limiting to just studies of persons with dementia severely limits the applicability of this review to the world that we work in. In general there is too much emphasis on diagnosis when this is an imperfect process at best. There are individuals with severe levels of MCI or MCI-R, for example, who suffer in multiple ways due to their disability. We need to identify these patients (Veterans) and "treat" them as well and often we do not address impairment until it is severe enough to be able to classify as "probable dementia."	The scope of this review was limited to dementia. However, we agree that Veterans with cognitive impairment non-dementia (CIND) are an important patient population. The best practices for assessment and management of patients with CIND may warrant a separate evidence review.
2	No.	Noted.

Reviewer	Comment	Response
3	Yes. Not a systematic bias but an odd ignoring of certain literature with over emphasis on other papers.	Noted. We have responded to the specific points raised in Question 4.
4	No - However, some who are strong advocates for screening might complain that you have only looked for adverse effects of screening, and have not considered adverse effects of missed diagnosis (e.g, worsened chronic disease control). To my knowledge these studies have not been done, so I personally do not believe this is a strong criticism.	Although identifying the consequences of falsely negative results from brief mental status tests was not within the scope of Key Question 2, we agree that the effects of case-finding and missed diagnoses on the improvement or worsening of patient outcomes are important considerations.
5	No bias was visible in this review.	Noted.
Question 3. Are there any studies on dementia signs and symptoms, or on the cognitive measures of interest to the VA, that we have overlooked?		
1	Yes. My only concern here is again the filter that was used, as a general concern about the completeness of your literature search. Some studies have used tools that estimate a high probability of dementia based on prior validation studies with that instrument. You need to be careful to include studies that also have this approximation even if subjects were not even ostensibly diagnosed. You did this when including Crooks' study, a study I know well and these patients were not "diagnosed" but had an approximation of diagnosis applied. I conducted a study of memory impairment in Veterans and concern about impairment as a strong predictor where there was a high probability of dementia that was dictated by conservative cut-points from prior validation research. This should have been considered for inclusion (perhaps it was) especially because it was done within the VA.	This was designed to be a review case-finding tools rather than tools that are predictive of future dementia. Re: the Crooks paper - we did not include this as a primary study - it was included as part of the Mitchell review and we did, in our revisions, clarify the weaknesses of individual studies from the Mitchell review including Crooks (we agree they did not use a gold standard for dementia assessment).
2	None that I'm aware of specifically.	Noted.
3	Yes. I think you gloss over a sizable literature on informant reported memory problems in favor of focusing on a patient's own subjective complaints. It has been suggested repeatedly that subjective memory complaints on the part of the patient are frequently associated with anxiety and depression while the reports of informants (and a literature suggesting that spouses are the best informants) are more related to actual cognitive decline yet you avoid discussion of the problems with an individual's complaints and ignore informant complaints almost completely. The MMSE has many problems and is just about as insensitive as the mini-Cog and short Blessed (BOMC), but it has the largest body of literature by far and is most familiar to actual providers.	We've added some more discussion points about informant reported memory problems. We had not included some studies (eg - Archer 2007) because they fell outside inclusion criteria (assessing MCI in this case). Many studies investigate association with future cognitive decline. However, we agree it is important to clarify some of the potential weaknesses of patient reported complaints and acknowledge the potential role of informant report. We've added the Jorm 1997 review to our dicussion, though it really looks at the value of longer informant questionnaires which is slightly different from "signs/sx" (which a very brief elicitation of SMC might approximate). We added Carr 2000 as well. Re: MMSE - it was simply outside of our review's scope - the scope of our report reflects the Dementia Steering Committee's interests - they were interested in literature about six commonly used alternatives to MMSE.

Reviewer	Comment	Response
4	See comment above concerning dementia warning signs considered in combination, rather than in isolation. Also – there are other warning signs type instruments – see for example: Galvin JE, et al, The AD8, a brief informant interview to detect dementia,Neurology 2005:65:559-564. See also the following article that may help to illuminate the role of informant information and its usefulness in guiding diagnosis of dementia. Informant ratings of cognitive decline in old age: validation against change on cognitive tests over 7 to 8 years. Jorm AF. Christensen H. Korten AE. Jacomb PA. Henderson AS. Psychological Medicine. 30(4):981-5, 2000 Jul. Validation analysis of informant's ratings of cognitive function in African Americans and Nigerians. Shen J. Gao S. Unverzagt FW. Ogunniyi A. Baiyewu O. Gureje O. Hendrie HC. Hall KS. International Journal of Geriatric Psychiatry. 21(7):618-25, 2006 Jul.	Agree - we've added more information re: informant report including the articles you mention and others we found. We mentioned the Shen study in the discussion and future studies section. The warning signs in combination is an interesting point and has not been well-studied - we've added it to suggestions for future studies.
5	Yes. One important document to consider (if not already considered) is the Alzheimer's association 10 warning signs for dementia. I find that these warning signs, which were recently updated after significant input, are worth sincere evaluation as dementia warning signs.	The suggested document was included in the draft report.
Question 4. Please write additional suggestions or comments below. If applicable, please indicate the page and line numbers from the draft report.		
1	I am concerned about restrictive (and traditional) intentions of this work: 1) find ways to identify and diagnose dementia patients or those at high risk; 2) do this in a brief "cost-effective" way; and 3) find signs and symptoms because screening poses too much of a burden. Medical providers are drilled on signs and symptoms of CHF, for example, and we have Review of System questions to identify potential risk. Practitioners are often uncomfortable about screening or ruling in or out dementia. Your review suggests that warning signs are not helpful – at least there is little evidence thus far. Perhaps more attention needs to be put on how to teach providers how to sensitively and comfortably discuss cognitive concerns. Your review and others like it suggest that memory complaints might not predict dementia (though I believe that my study did predict it) and therefore they do not need to be addressed. I think you should include a discussion about addressing patient complaints regarding memory because of the inherent need to do that as an effective and sensitive practitioner regardless of the underlying reason for the complaint. A normal screen can be very reassuring and may address significant anxiety. With all that said, this is a very professional, well organized and thoughtful review. This has real value to VA and its Veterans.	These are well-considered and well-put concerns. Much of this debate is beyond the scope of this review. Because the issue of screening for dementia (as distinguished from case-finding) has been widely debated and there are already excellent publications and reviews outlining both sides of this debate, we have tried to make clear this review does not address the relative value of widespread screening for dementia. The points re: training providers how to sensitively and comprehensively address memory concerns, and the reassurance that a normal screen can offer a patient are well-taken. We have inserted some suggested future areas of study along these lines.

Reviewer	Comment	Response
2	Please see attached document with comments in Track Changes. This is an important, well-written report. But we need the methods/results/discussion/ conclusions/ clearly presented in non-technical terms so that non-expert policy-makers will understand the recommendations for further action. Don't "hide" the results and recommendations within technical terms.	Noted, and we've modified the language.
2	Background in body and Exec Summary, Lines 3-7: these data come from VA Allocation Resource Center (ARC) data, which are on the VA Intranet only and are not public. Citation 1 references the ARC website on the VA Intranet. It has not yet been decided what parts of the DSC Report can be made public. An alternative would be to quote the 2004 VA dementia projection from the VHA Office of the Assistant Deputy Under Secretary for Health for Policy and Planning, also posted on the internet at www1.va.gov/vhareorg/reports.htm. There is also a 2003 publication by VA HERC that gives some dementia cost data from 1999 (Yu W, et al, 2003). These public data are mentioned on p. 1-2 of the DSC report.	We have replaced this paragraph using data from the suggested references.
2	Page v line 27; also page 17 line 12: sterognosis and graphesthesia: use layperson terms in addition to the technical terms, for non-expert readers	We have defined these terms in the text, and added them to the Glossary.
2	Methods, page 3, line 8: citation 13 is an internal memo, I'm not sure it is appropriate to cite it here.	We agree and have deleted reference 13.
2	Page 4, lines 5-6: Can you say more about the decision to exclude studies on signs and symptoms that predict future incident dementia? (e-mail discussion copied here) ...A little more explicit would be great. If you could amplify in the report just a bit, like your sentence below that I highlighted in yellow, that would be helpful. It may be just a matter of adding the "layperson" wording, or repeating it in the Study Selection section as you did in the Background section.	This was really designed as a review of case-finding tools rather than a review of tools that are predictive of future dementia. Our review does not address the latter issue and I hope this is clear to the reading audience. Given the lack of convincing evidence re: screening for dementia and the lack of clear consensus around this issue, we had thought that a review of methods for predicting dementia would be less applicable. We tried to frame this issue in the introduction, but again please let me know if we should be more explicit.
2	Page 16, line 9: Word(s) missing? Check sentence: "...questionnaires were completed (not?) by caregivers of demented participants, but by the non-demented participants themselves (22)."	We have clarified this sentence to read as follows: "Results from the third study are more difficult to interpret because sleep disturbance questionnaires were completed by caregivers of demented participants, whereas non-demented participants in the control group completed the questionnaire themselves (22)."
2	Page 32, lines 11 and 16: "cognitive screening instruments" - Define "screening" as used here. To me, screening means evaluation of asymptomatic individuals; you may not mean that here. Or reword, e.g. "which brief mental status instruments were used..."	We have reworded the phrase as suggested.

Reviewer	Comment	Response
3	My concern is that the literature is misrepresented. I do not think some of these papers can be correctly interpreted in the way you are trying to do here (particularly the behavioral signs and sleep sections). For the performance of the screening tests themselves, strong bias is exerted in the way the studies were done. To ignore these methodological issues is to present data that is much skewed from the actual performance of these instruments in a primary care setting.	Noted - specific responses detailed below.
3	This is a difficult literature. Results from the VA primary care population that would be most directly relevant aren't available because of the time it has taken us to get things written up. Still, I think you need to at least factor in the likely bias introduced by the published papers. I have detailed comments below. The performance of these tests is very different than that you present, supporting the idea that bias was introduced in the study design.	Noted - specific responses detailed below. Also, we had completed extensive quality evals of studies and will include these in our appendices.
3	I wouldn't put too much emphasis on that one meta analysis (ref 15). The individual studies included in the cross sectional analyses had some pretty iffy ways that dementia was called. For instance, the mmse is so insensitive that by the time a patient has crossed their threshold, the impairment was too far advanced for subjective memory complaints to be relevant to a case finding scenario in the clinic. I would have similar concerns about the use of a telephone screener to establish a diagnosis of dementia.	We re-wrote this section and tried to highlight more the deficiencies of some of the included studies.
3	I think the use of ADAS-Cog for diagnosis of dementia is questionable also. There may be some who would support it, but it is hardly a 'gold standard' as you call it (ref 35).	Noted and manuscript updated
3	Ref 36 had a good cognitive evaluation but this study examined different definitions of mci. When one of the criteria for diagnosis is subjective memory complaints, it does not seem valid to then look at how subjective memory complaints do as an indicator of the diagnosis.	Noted in manuscript
3	Ref 37 used an informant interview as their 'gold standard'. This was done with 121 subjects who screened positive and 35 who screened negative on a brief screen. While I believe that informant information is more valuable in this area that patient subjective memory complaint, I don't think this is a valid diagnostic method. This study establishes the level of correlation between patient complaint and informant complaint, nothing more.	Noted in manuscript

Reviewer	Comment	Response
3	As for reference 40, again the telephone screener was the primary means of evaluation. I suspect many, many milder cases were missed. Additionally, the subjective memory complaint was based solely on the question "do you have severe memory impairment?" as reported on a mailed questionnaire. I don't know of validation of this method, and it isn't really a straightforward inquiry into subjective memory problems. You go on to say that this method could classify healthy individuals as demented thus exaggerating reported specificity. The risk is actually higher the other way, I'd posit.	Noted and this study was de-emphasized
3	How did ref 43 get included? It doesn't compare prevalence rates of subjective memory impairment in those with and without dementia.	True - it compares SMC in patients with various levels of cognitive performance, but does not establish dementia diagnoses -- will exclude
3	Starting on line 15 of page 14 under the heading "Detailed Description" the use of MMSE is incorrect here. The Cache County study used an expanded and considerably more sensitive modified mini mental state exam or 3MS. I also believe you are misrepresenting/misunderstanding this study. While 61% of the demented subjects had behavioral disturbances in the past month (hardly low symptom prevalence as you state), this study documented behavioral problems in dementia and was not intended to address the use of behavioral symptoms in screening.	Thanks for clarifying - we will amend the report accordingly re: the modified MMSE. We hadn't intended to represent this as a study of use of behavorial sx for screening. Rather, the study was intended to test the hypothesis that in a community-based sample, demented patients would have more neuropsychiatric sx than non-demented patients. This may be useful for case-finding - ie - which sx should prompt a primary care physician to assess for dementia. Many clinicians may not think to assess one for dementia if presenting with sx such as apathy. Re: sx prevalence, we had meant to say neuropsych sx were common, but the prevalence of any one sx was relatively low. We will clarify in the report.
3	Page 15, line 17/18, do you mean the accuracy of dementia documentation in medical records of persons with and without dementia or with and without depression? If it is the former there is a whole literature that you ignore that concerns the documentation of a dementia diagnosis in the primary care setting.	We meant the latter - we do reference the former issue in the background section.
3	For reference 20, given that the average mmse of those they consider undiagnosed dementia was 17, I suspect they missed the mild cases. A MMSE of 17 isn't on the borderline of diagnosis. I'm also not convinced of their depression diagnoses, but setting aside the inadequacies of the diagnoses, I wouldn't jump to the conclusion that "some persons with mild dementia may have been misclassified as being depressed" (page 15, line 22/23). I don't think they caught those with mild dementia but the relationship between depression and dementia is nuanced and complex for even experienced geriatricians and geriatric psychiatrists.	Agreed - last statement was cut.

Reviewer	Comment	Response
3	You cannot look at a sample of Parkinson's disease patients and say anything about sleep and dementia except in the specific case of sleep and dementia in Parkinson's disease (top page 16).	We agree - our statement is meant to refer only to this study, but we will make clearer in the text that the results don't apply to patients with other types of dementia.
3	Reference 21 is incomplete.	It had originally been published online - we fixed the ref
3	Reference 28 is too old for me to access it electronically, but if frontal release signs were present in over half the participants, this is a severely demented group of subjects and not relevant to your topic.	Mean MMSE scores in the demented group were about 21 and suggested this was not a group with severe dementia. I agree this was not representative of a screen-diagnosed group of patients, but the findings may be relevant to case-finding discussions. Also interesting is that 9% of controls had release signs.
3	Reference 27 is also available to me only in abstract, but the abstract says 'with the exceptions of impaired vibration sense, loss of upward gaze, and bradykinesia, all signs were associated with the neurodegenerative syndromes and stroke' which seems different than what you report.	The abstract is slightly misleading - many of the signs were associated with stroke and/or Parkinsons, not dementia. The signs we mentioned were the ones associated with dementia. The study has numerous flaws in any case.
3	Reference 30 correlates an individual's assessment of their driving ability with that of an experienced neurologist and a driving instructor. It doesn't address use of driving skills as a screen for cognitive ability. Again, I think you misrepresent the literature.	It technically fit our inclusion criteria - I agree it's not a very useful study and we amended the paragraph to clarify the focus of the paper (was in the table, but I agree could have been clearer).
3	For reference 48, can you really report sensitivities of the Short Blessed (BOMC) of 100%? The 'field diagnoses' of dementia are suspect. Those felt to be demented were invited to Duke for a real evaluation. There were seven of these. The Short Blessed is brief and easily memorized by clinicians and requires no props. It is not, however, a test with a sensitivity of 100%.	We agree that sensitivity may be overestimated because 13 subjects had a documented history of dementia, among the 26 identified as probably demented by field diagnosis. We have noted this limitation in Table 4 and in the text of the Results, in the BOMC section. Because the field methods used DSM criteria as a guide, the study meets our criteria for inclusion. However, we have replaced the results from this study in Table 5 with the results from Stuss 1996, a memory clinic sample that did not include patients who had been previously diagnosed with dementia.
3	Ref 49, again too old to pull up, seems to focus on severe dementia. Do your reported sens/spec come from a severely demented subgroup?	The sample in this study (Stuss 1996) consisted of patients who had been referred to a memory clinic for possible dementia, including some who on successive evaluations turned out not to have dementia. The final diagnosis was determined subsequent to the BOMC test. Although the results of the referral sample may not be applicable to primary care populations, the patients did not have a history of dementia. It does not appear that the sens/spec results were weighted by a more severely demented subgroup.

Reviewer	Comment	Response
3	Why is there a paragraph on informant reports of memory loss on pg 28? Again, you ignore a whole literature on this topic to focus on this paper?	We agree that this paragraph was not relevant to this section, and have removed it accordingly.
3	I really think that you ought to stress more that the six measures you look at are chosen by an 'expert panel' for various clinical reasons, not for rigorous scientific reasons. This was a series of phone conferences with people who were tasked with finding alternatives to the MMSE.	We agree, and have added text to the Methods section to indicate that the six measures reviewed were based on several clinical criteria.
3	The mini-Cog was not actually administered in MoVIES (ref 56) or in the University of Washington ADC (ref 64). Items incorporated in testing actually done were pulled to create a miniCog score. In actual practice, the sensitivities of the miniCog are considerably lower. Again, a brief test, easily remembered by clinicians and requiring no props except a paper and pencil, but not that sensitive.	We have added text to the Results to clarify that the Mini-Cog results from these studies were derived from components of longer tests, and that these results may not be directly comparable to the use of the Mini-Cog by itself in practice.
3	The diagnosis of dementia in ref 54 was based purely on informant interview, no cognitive testing. Not sufficient. Also, half of this ADC population was made up of non English speakers. How many of the informants spoke sufficient English to give a good history?	The study notes that non-English speakers were administered cognitive tests by foreign-born native speakers who were also fluent in English, although it is not explicitly specified that non-English speaking informants were interviewed likewise by an interpreter. However, we agree that this study does not meet criteria for the use of a full reference standard such as DSM-IV, because diagnostic workup for dementia was based on positive informant history, regardless of other patient evaluations that were conducted. We thank you for bringing this to our attention. We have excluded this study from our review.
3	It doesn't appear that investigators actually did any cognitive testing to arrive at a dementia diagnosis for the SLUMS study (ref 61). You comment on the inclusion of MCI in their population. They did not diagnose MCI (and could not without some testing) but rather a "mild neurocognitive disorder." There isn't such a diagnosis in DSM IV that I know of.	In Tariq 2006, the Methods state that each participant was evaluated during a routine clinic visit, a history was obtained from corroborating sources, a complete physical and mental status exam was performed, and lab findings were reviewed. Based on these data (which included a mental status exam) the investigators appear to have used, and state that they used, DSM-IV criteria to diagnose dementia. We concur that MCI is not a DSM-IV diagnosis; individuals with MCI fit neither criteria for normal or demented. Because different studies dealt differently with this subgroup, and some excluded them from analyses altogether, we noted how their inclusion could affect results.

Reviewer	Comment	Response
3	Page 32 under the heading of "Qualitative studies of the GPCOG and the Mini-Cog". This is a purely volunteer sample of clinicians. For a questionnaire with a 25% response rate (less actually since that 25% seems to be those who completed any part of the survey and there were the most responses familiar with the MMSE), there were 88% of respondents unfamiliar with the GPCOG and 75% unfamiliar with the Mini-Cog. This suggests giving weight to the 35 practitioners who said they had heard of the GPCOG and the 73 who said they'd heard of the miniCog.	We agree, and have removed the qualitative studies. We contacted the author of the IPA survey, who clarified that respondents rated the perceived quality on those tests with which they had some familiarity. Consequently the results for the GPCOG and Mini-Cog are based on a small, potentially biased sample. The 2nd qualitative study examined 2 of the 6 VA measures and ranked them equally high. Therefore we have noted in Future Research Recommendations that similar surveys be conducted among VA providers for input on the use of the 6 measures in practice.
3	As for the acceptability of screening, please note that you are basing your section primarily on one investigator using samples very different from the VA where we have had much better response from the Veterans concerning cognitive screening.	We have made note in the Results for KQ3 that 2 of the 3 studies were led by the same investigator. We have also removed some of the details about the studies and kept the more general findings, because these studies do not provide direct evidence about the 6 cognitive measures used in VA.
5	I found this report very informative overall and it does justice to this large body of literature pertaining to this important topic.	Noted.

APPENDIX E. GLOSSARY

Agnosia: Failure to recognize or identify objects despite intact sensory function.

Alzheimer's disease (AD): A disease usually characterized by loss of memory, especially for learning new information, reflecting deterioration in the functioning of the medial temporal lobe and hippocampus areas of the brain. Later in the illness, other higher functions of the cerebral cortex become affected: these include language, praxis (putting theoretical knowledge into practice) and executive function (involved in processes such as planning, abstract thinking, rule acquisition, initiating appropriate actions and inhibiting inappropriate actions, and selecting relevant sensory information). Behavioral and psychiatric disturbances are also seen, which include depression, apathy, agitation, disinhibition, psychosis (delusions and hallucinations), wandering, aggression, incontinence and altered eating habits.

Aphasia: Deterioration in language skills, such as word-finding difficulties, reduction in output, loss of fluency, and poor comprehension.

Apraxia: Total or partial loss of the ability to perform coordinated movements or manipulate objects in the absence of motor or sensory impairment.

Case-finding: The strategy of identifying a new occurrence of disease among patients selected on the presence of risk factors, signs, or symptoms.

Cohort: A group of persons with a common characteristic or set of characteristics. Typically, the group is followed for a specified period of time to determine the incidence of a disorder or complication of an established disorder (prognosis).

Cohort Study: (Cohort Analytic Study): Prospective investigation of the factors that might cause a disorder in which a cohort of individuals who do not have evidence of an outcome of interest but who are exposed to the putative cause are compared with a concurrent cohort who are also free of the outcome but not exposed to the putative cause. Both cohorts are then followed to compare the incidence of the outcome of interest. Used for Prospective Study.

Dementia: The development of multiple cognitive deficits that include memory impairment and at least 1 of the following cognitive disturbances: agnosia, aphasia, apraxia, or a disturbance in executive functioning. Deficits must be severe enough to cause significant decline in social or occupational functioning and must represent a decline from previous baseline functioning.

Dementia with Lewy bodies (DLB): One of the most common types of progressive dementia and shares characteristics with both Alzheimer's and Parkinson's diseases. Its central feature is progressive cognitive decline, combined with three additional defining features: pronounced fluctuations in alertness and attention, such as frequent drowsiness, lethargy, lengthy periods of time spent staring into space or disorganised speech; recurrent visual hallucinations; and parkinsonian motor symptoms, such as rigidity and the loss of spontaneous movement. The symptoms of DLB are caused by the build-up of Lewy bodies (protein deposits found in nerve cells) in areas of the brain that control particular aspects of memory and motor control.

Executive Functioning: The ability to think abstractly and to plan, initiate, sequence, monitor, and stop complex behavior.

Graphesthesia: The ability to recognize a number or letter written on the skin by the sensation of touch.

Heterogeneity: A term used to illustrate the variability or differences between studies in the estimates of effects.

Incidence: The number of new cases of a disease per population over a given time period. Incidence measures how frequently a new case of disease occurs, as opposed to prevalence, which conveys how widespread a disease is in a population.

Inter-rater reliability: A measure of the extent to which multiple raters or judges agree when providing a rating, scoring, or assessment.

Key questions: Questions posed by the advisory panel that are used to guide the identification and interrogation of the evidence base relevant to the topic of the guideline.

Likelihood Ratio: For a screening or diagnostic test (including clinical signs or symptoms), expresses the relative likelihood that a given test result would be expected in a participant with (as opposed to one without) a disorder of interest.

Mild Cognitive Impairment (MCI): Presence of a memory complaint, preferably corroborated by an informant, objective memory impairment, and normal general cognitive function. Activities of daily living are intact and the patient does not meet clinical criteria for dementia.

Neuropsychiatric symptoms: Symptoms which generally fall within one of three symptom clusters: agitation, psychosis and mood disorders. Symptoms of agitation often include aggressiveness or irritability. Symptoms of psychosis are hallucinations – auditory or visual, and delusions. Mood disorders would include depression and anxiety.

Odds ratio: A ratio of the odds of having the disease of interest in a group with a particular exposure, symptom, or characteristic of interest, to the odds of disease in a group that does not have the exposure /symptom / characteristic. An odds ratio of 1 indicates that the disease is equally likely to occur in both groups. On odds ratio of 4 indicates that the disease is 4 times more likely to be present in the group that has the symptom or characteristic of interest, compared with the group that does not have this symptom.

Predictive Value (PPV): Positive Predictive Value – the proportion of people with a positive test who have the disease; Negative Predictive Value – proportion of people with a negative test who are free of disease.

Prevalence: The total number of cases of the disease in the population at a given time, expressed as a proportion in which the number of cases is the numerator and the population at risk is the denominator.

Release signs: Primitive reflexes that are normally present in infants, including the suck, snout, palmomental, and grasp reflexes. They are seen in disorders that affect the frontal lobes, such as dementias, metabolic encephalopathies, closed head trauma, and hydrocephalus.

Remote/over-learned memory: The ability to remember people or events from the distant past, or over-learned information such as the days in the week or one's birthday.

Registration/recall: The processing of received information, and the retrieval of the information in response to a cue for use in a process or activity.

Screening: A strategy used in a population to detect a disease in individuals without signs or symptoms of that disease. Universal screening involves screening of all individuals in a certain category, for example, all persons age 65 and older.

Screening test: A brief instrument used to determine the likelihood of whether a disease may be present, and whether more comprehensive diagnostic testing may be needed. The predictive value of a screening test is influenced by the prevalence of the disease in the population. In universal screening, a screening test would be administered to all persons in a certain category (e.g. age 65+). In a case-finding approach, the screening test would be selectively administered when a patient has risk factors or presents with signs or symptoms of the disease.

Sensitivity: The proportion of people who truly have a designated disorder who are so identified by the test. The test may consist of, or include, clinical observations. The proportion of truly diseased persons in the screened population who are identified as diseased by the screening test—that is, the true-positive rate.

Specificity: The proportion of people who are truly free of a designated disorder who are so identified by the test. The test may consist of, or include, clinical observations. The proportion of truly nondiseased persons who are identified as such by the screening test—that is, the true-negative rate.

Stereognosis: The ability to perceive the form of an object by using the sense of touch.

Sources

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition, Text Revision. Washington, DC: American Psychiatric Association, 2000.

Guyatt G, Rennie D. User's Guides to the Medical Literature: A manual for evidence-based clinical practice. Chicago: AMA Press, 2002.

Holsinger T, Deveau J, Boustani M, Williams JW, Jr. Does this patient have dementia? JAMA 2007;297(21):2391-404.

National Institute for Health and Clinical Excellence (NICE). Dementia: A NICE–SCIE Guideline on supporting people with dementia and their carers in health and social care. In: National Clinical Practice Guideline Number 42: The British Psychological Society and Gaskell; 2007.

Petersen RC, Stevens JC, Ganguli M, Tangalos EG, Cummings JL, DeKosky ST. Practice parameter: early detection of dementia: mild cognitive impairment (an evidence-based review). Report of the Quality Standards Subcommittee of the American Academy of Neurology. Neurology 2001;56(9):1133-42.

Walker, HK. The suck, snout, palmomental, and grasp reflexes. *Clinical Methods – the History, Physical, and Laboratory Examinations.* Third Edition. Butterworth Publishers, 1990.

West S, King V, Carey TS, et al. Systems to rate the strength of scientific evidence: Evidence Report/Technology Assessment, Number 47. Rockville, MD: Agency for Healthcare Research and Quality. AHRQ Publication No. 02-E016, April 2002.