International Psychogeriatrics (2014), 26:4, 555–563 © International Psychogeriatric Association 2013
doi:10.1017/S1041610213002329

REVIEW

A meta-analysis of the accuracy of the Addenbrooke's Cognitive Examination (ACE) and the Addenbrooke's Cognitive Examination-Revised (ACE-R) in the detection of dementia

Andrew J. Larner[1] and Alex J. Mitchell[2,3]

[1]*Cognitive Function Clinic, Walton Centre for Neurology and Neurosurgery, Liverpool, UK*
[2]*Department of Psycho-oncology, Leicestershire Partnership NHS Trust, Leicester, UK*
[3]*Department of Cancer and Molecular Medicine, University of Leicester, Leicester, UK*

ABSTRACT

Background: The Addenbrooke's Cognitive Examination (ACE) and its Revised version (ACE-R) are relatively new screening tools for cognitive impairment that may improve upon the well-known Mini-Mental State Examination (MMSE) and other brief batteries. We systematically reviewed diagnostic accuracy studies of ACE and ACE-R.

Methods: Published studies comparing ACE, ACE-R and MMSE were comprehensively sought and critically appraised. A meta-analysis of suitable studies was conducted.

Results: Of 61 possible publications identified, meta-analysis of qualifying studies encompassed 5 for ACE (1,090 participants) and 5 for ACE-R (1156 participants); of these, 9 made direct comparisons with the MMSE. Sensitivity and specificity of the ACE were 96.9% (95% CI = 92.7% to 99.4%) and 77.4% (95% CI = 58.3% to 91.8%); and for the ACE-R were 95.7% (95% CI = 92.2% to 98.2%) and 87.5% (95% CI = 63.8% to 99.4%). In a modest prevalence setting, such as primary care or general hospital settings where the prevalence of dementia may be approximately 25%, overall accuracy of the ACE (0.823) was inferior to ACE-R (0.895) and MMSE (0.882). In high prevalence settings such as memory clinics where the prevalence of dementia may be 50% or higher, overall accuracy again favored ACE-R (0.916) over ACE (0.872) and MMSE (0.895).

Conclusions: The ACE-R has somewhat superior diagnostic accuracy to the MMSE while the ACE appears to have inferior accuracy. The ACE-R is recommended in both modest and high prevalence settings. Accuracy of newer versions of the ACE remain to be determined.

Keywords: Addenbrooke's Cognitive Examination, meta-analysis, Mini-Mental State Examination; sensitivity and specificity

Introduction

The Addenbrooke's Cognitive Examination (ACE) is a brief cognitive test battery which was first published in 2000 (Mathuranath *et al.*, 2000) as a theoretically motivated improvement over the Mini-Mental State Examination (MMSE) (Folstein *et al.*, 1975). It aimed to address some of the recognized neuropsychological deficiencies of the MMSE by adding more tests of memory and visuospatial abilities and introducing tests specifically examining executive function. Although the ACE was widely and rapidly accepted in clinical practice, certain weaknesses were identified, prompting the development of the Addenbrooke's Cognitive Examination-Revised (ACE-R) in 2006 (Mioshi *et al.*, 2006). The item content of ACE and ACE-R is compared with the MMSE in Table 1.

There have been a significant number of publications examining the accuracy of the ACE and ACE-R in various patient populations using both the original instruments and translations into a variety of languages (Davies and Larner, 2013). Results from a National Dementia Register suggested that ACE-R items improved MMSE estimates of cognitive ability by 16% (Law *et al.*, 2013). A systematic review of ACE/ACE-R studies identified nine of 45 studies as suitable for review, excluding translated versions, and concluded that ACE/ACE-R was capable of differentiating between those with and without cognitive impairment,

Correspondence should be addressed to: Dr Larner, Cognitive Function Clinic, Walton Centre for Neurology and Neurosurgery, Liverpool, UK. Email: a.larner@thewaltoncentre.nhs.uk. Received 21 Aug 2013; revision requested 13 Sep 2013; revised version received 2 Nov 2013; accepted 6 Nov 2013. First published online 5 December 2013.

Table 1. Item content of ACE, ACE-R, and MMSE

	ACE	ACE-R	MMSE
Orientation: Time	5	5	5
Orientation: Place	5	5	5
Registration	3	3	3
Attention/Concentration (serial 7s, DLROW)	5 (best performed task)	5 (best performed task)	5
Memory: Recall	3	3	3
Memory: Anterograde memory	28	19	
Memory: Retrograde memory	4	4	
Verbal fluency: Letters and animals	14	14	
Language: Naming	12	12	2
Language: Comprehension	8	8	4
Language: Repetition	5	4	1
Language: Reading	2	1	
Language: Writing	1	1	1
Visuospatial abilities: Intersecting pentagons	1	1	1
Visuospatial abilities: Wire (Necker) cube	1	2	
Visuospatial abilities: Clock drawing	3	5	
Perceptual abilities: Dot counting	–	4	
Perceptual abilities: Fragmented letters	–	4	
Total score	100	100	30

but the evidence base on distinguishing dementia subtypes and mild cognitive impairment (MCI) was lacking (Crawford *et al.*, 2012). However, to date no meta-analysis of ACE/ACE-R studies has been published to our knowledge.

The aim of this study was to identify and quantify robust studies reporting the diagnostic accuracy of ACE and ACE-R in relation to dementia versus non-demented subjects.

Methods

Methods were similar to those used previously for meta-analysis of MMSE dementia studies (Mitchell, 2009, 2013). Guidelines for the performance and reporting of meta-analyses in the PRISMA statement (Liberati *et al.*, 2009; Moher *et al.*, 2009) were observed.

Search criteria

A systematic literature search was undertaken using the following abstract databases: Medline 1966-May 2013, PsycINFO 1887-May 2013, and Embase 1980-May 2013. Five full text collections (Science Direct, Ingenta Select, Ovid Full Text, Blackwell Online and Wiley Interscience) were searched. The abstract database Web of Knowledge (4.0, ISI) was searched. The following search terms were used: "Screen* or test or instrument or measure or tool or diagnose* or Addenbrooke's Cognitive Examination or ACE or ACE-R" and "dementia or Alzheimer* or cognitive" and "sensitivity and specificity or accuracy or cut-off or receiver operator or ROC or Youden."

Inclusion criteria

Studies were included that examined diagnostic validity of ACE or ACE-R in comparison to a validated diagnostic standard of either dementia or MCI (see Tables 2 and 3 for details).

Exclusion criteria

Studies were excluded that did not present adequate data for inclusion in the analysis. Based on previous sample size calculations, studies that recruited less than 160 patients (80 per group) were excluded to avoid studies with unreliable false negative or false positive results (Mitchell, 2009). No exclusion was made solely on the basis of language of ACE/ACE-R publication (Crawford *et al.*, 2012), as these instruments have been translated into many languages (Davies and Larner, 2013).

Analysis

Studies identified were subjected to critical appraisal and a meta-analysis of suitable studies was conducted, to give overall test accuracy, sensitivity, specificity, combined Youden score, positive and negative predictive values (PPV, NPV), positive and negative likelihood ratios (LR+, LR-), and positive and negative clinical utility index (CUI+, CUI-) (Mitchell, 2009). The interpretation of the clinical utility index is 0.93–1.00 near perfect value; 0.81–0.92 excellent; 0.64–0.80 good; 0.49–0.63 adequate; 0.36–0.48 poor; and < 0.36 very poor. Publication bias was tested by Harbord method (Harbord *et al.*, 2006). Comparative accuracy was tested by conducting a relative risk comparison of

Table 2. Diagnostic validity of ACE in diagnosing dementia

REFERENCE	REFERENCE STANDARD	SAMPLE DESCRIPTION	CUT-OFF	SENSITIVITY (%)	SPECIFICITY (%)	PPV (%)	NPV (%)	AUC ROC	PREVALENCE
Mathuranath *et al.*, 2000	DSM-IV, NINCDS-ADRDA, NINDS-AIREN, Neary FTD	$N = 266$, 115 with dementia, normals 127, non-dementia 24	88/100	93	71	71	93	0.91	0.43
Garcia-Caballero *et al.*, 2006	DSM-IV, NINCDS-ADRDA, NINDS-AIREN, Neary FTD	$N = 167$, 70 with dementia, normals 72, non-dementia 25	68/100	91	86	82	93	0.96	0.42
Larner 2007	DSM IV	$N = 285$, 140 with dementia, 145 non-dementia clinic referrals	88/100	100	43	63	100	0.93	0.49
Stokholm *et al.*, 2009	NINCDS-ADRDA	$N = 171$, 78 with dementia, normals 63, depression 30	85/100	99	94	95	98	0.99	0.46
Yoshida *et al.*, 2011	DSM-IV, NINCDS-ADRDA, NINDS-AIREN, Neary FTD, McKeith DLB 2000, Petersen MCI	$N = 201$, 126 with dementia, normals 62, MCI 13	80/100	98	87	93	97	N/A	0.63

Table 3. Diagnostic validity of ACE-R in diagnosing dementia

REFERENCE	REFERENCE STANDARD	SAMPLE DESCRIPTION	CUT-OFF	SENSITIVITY (%)	SPECIFICITY (%)	PPV (%)	NPV (%)	AUC ROC	PREVALENCE
Mioshi *et al.*, 2006	NINCDS-ADRDA, Neary FTD, McKeith DLB 2000, Petersen MCI	$N = 241$, 142 with dementia, normals 63, MCI 36	88/100	94	89	92	91	N/A	0.59
Alexopoulos *et al.*, 2010	NINCDS-ADRDA, Lund/Manchester FTD, Winblad MCI	$N = 229$, 78 with dementia, normals 76, MCI 75	82/100	92	96	94	99	0.99	0.34
Yoshida *et al.*, 2012	DSM-IV, NINCDS-ADRDA, NINDS-AIREN, Neary FTD, McKeith DLB 2000, Petersen MCI	$N = 242$, 130 with dementia, normals 73, MCI 39	82/100	99	99	99	99	0.99	0.54
Dos Santos Kawata *et al.*, 2012	DSM-III, NINCDS-ADRDA, Neary FTD, McKeith DLB 1996, Emre PDD	$N = 201$, 126 with dementia, normals 85	80/100	94	94	96	91	0.98	0.63
Larner, 2013a	DSM IV	$N = 243$, 84 with dementia, 159 non-dementia clinic referrals	88/100	99	44	48	99	0.94	0.35

pooled sensitivity and specificity and by comparing overall accuracy at equivalent prevalence rates of 25% and 50%.

Results

Search results

The literature search identified 61 possible papers on ACE and ACE-R in all (full list obtainable from corresponding author; see Figure 1 for PRISMA flowchart). Of these, 29 were reports of studies of the ACE, 13 using the English version and 16 using non-English versions. There were 32 reports of studies of the ACE-R, 17 using the English version and 15 using non-English versions. After application of exclusion criteria, ten studies were suitable for meta-analysis, five for ACE (Mathuranath *et al.*, 2000; Garcia-Caballero *et al.*, 2006; Larner, 2007; Stokholm *et al.*, 2009; Yoshida *et al.*, 2011; Table 2), and five for ACE-R (Mioshi *et al.*, 2006; Alexopoulos *et al.*, 2010; Dos Santos Kawata *et al.*, 2012; Yoshida *et al.*, 2012; Larner, 2013a; Table 3).

All the studies identified were from high prevalence specialist secondary care settings. No study was identified from a low prevalence primary care or community setting. Most were experimental studies, applying the test to groups of patients with known diagnoses. Pragmatic studies, examining all clinic attenders before diagnosis was established, and hence reflecting the idiom of clinical practice, were in the minority.

Of the studies that examined ACE/ACE-R accuracy in relation to MCI, none identified was deemed suitable according to specified inclusion/exclusion criteria.

Criterion (gold) standard

The most common reference standard for dementia diagnosis was the Diagnostic and Statistical Manual (DSM) in all studies (usually version IV), with additional criteria for dementia subtypes.

Cut-off threshold

The original studies of ACE and ACE-R cited two different cut-offs, 88/100 and 83/100 (Mathuranath *et al.*, 2000), and 88/100 and 82/100 (Mioshi *et al.*, 2006), respectively. These cut-offs have been used in some of the subsequent studies, but not all, as different authors have examined different cut-offs according to the educational profile of their cohort (e.g. Garcia-Caballero *et al.*, 2006) or selecting the cut-off corresponding to the best overall accuracy of the test (Larner, 2007, 2013a). For the purposes of the pooled analysis, where studies had used more than one cut-off point the most sensitive was selected.

The accuracy of ACE in identifying dementia versus subjects without dementia

Across the 5 included studies there were 529 cases of dementia out of a population of 1,090, a prevalence of 49%. There was no evidence of publication bias (Harbord bias = –8.23, 95% CI = –29.1 to 12.6, p = 0.37).

Pooling the raw data from these studies demonstrated that 512 out of 529 cases were correctly identified using the ACE, giving a pooled Se of 96.8%. On meta-analytic weighting this was corrected to 96.9% (95% CI = 92.7% to 99.4%). Non-cases (377) were correctly ruled out from a sample of 561 comparison subjects to give a pooled Sp of 67.2%. On meta-analysis, this was corrected to 77.4% (95% CI = 58.3% to 91.8%). Unadjusted the PPV was therefore 74.7% and the NPV 95.5%.

The accuracy of ACE-R in identifying dementia versus subjects without dementia

Across the 5 included studies there were 560 cases of dementia out of a population of 1,156, a prevalence of 48%. Harbord bias was not significant (0.097, 95% CI = −18.95 to 19.14, p = 0.99).

Pooling the raw data from these studies demonstrated that 514 out of 560 cases were correctly identified using the ACE-R, giving a pooled Se of 91.8%. This was adjusted on meta-analytic weighting to 95.7% (95% CI = 92.2% to 98.2%). Non-cases (383) were correctly ruled out from a sample of 596 comparison subjects to give a pooled Sp of 64.3%. This was corrected on meta-analysis to 87.5% (95% CI = 63.8% to 99.4%). Unadjusted the PPV was therefore 70.7% and the NPV 89.3%.

Comparison of ACE and ACE-R with MMSE at given cut-offs

Five studies compared the ACE with the MMSE. For the MMSE, Harbord bias was 7.18 (95% CI = –24.77 to 39.14, p = 0.59). At a cut-off of <27 the adjusted meta-analytic sensitivity was 89.9% (95% CI = 81.9% to 95.7%). Specificity was 81.7% (95% CI = 71.6% to 90.0%). A head-to-head relative risk analysis suggested that the ACE was more accurate than the MMSE in terms of sensitivity (RR_{sens} = 1.07, 95% CI = 1.00 to 1.14, $\chi^2 = 4.3$, p = 0.0383) but no different in terms of specificity (RR_{spec} = 0.92, 95% CI = 0.75 to 1.13, $\chi^2 = 0.62$, p = 0.429).

Four studies compared the ACE-R with the MMSE. For the MMSE Harbord bias was –20 (95% CI = –115.6 to 75.6, p = 0.55). At mixed

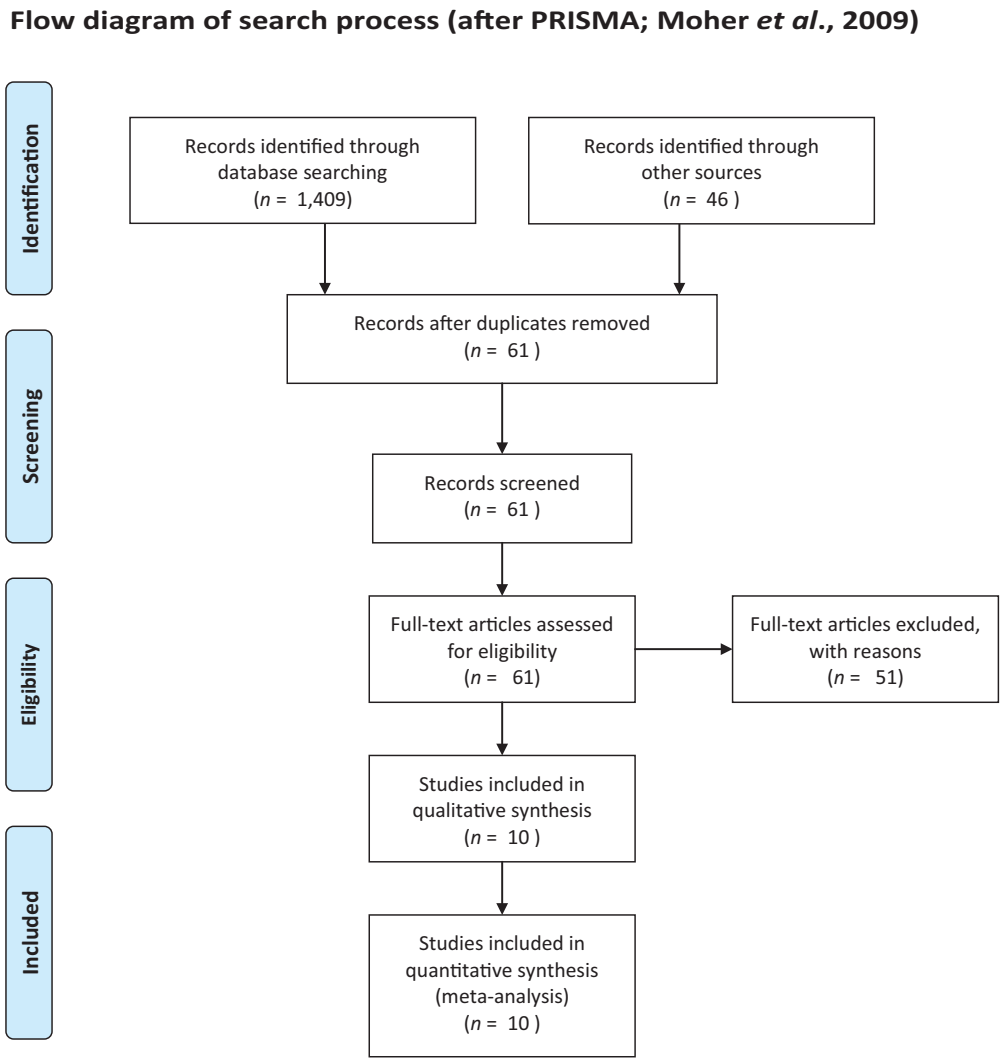

Figure 1. (Colour online) Flow diagram of search process (after PRISMA; Moher *et al.*, 2009).

cut-off points, the adjusted meta-analytic sensitivity was 94.0% (95% CI = 81.0% to 99.8%). Specificity was 92.1% (95% CI = 86.3% to 96.4%). A head-to-head relative risk analysis suggested that the ACE-R was no more accurate than the MMSE in terms of sensitivity ($RR_{sens} = 1.05$, 95% CI = 0.92 to 1.18, $\chi^2 = 0.51$, p = 0.470) and also no different in terms of specificity ($RR_{spec} = 0.87$, 95% CI = 0.64 to 1.18, $\chi^2 = 0.82$, p = 0.364).

Combining all nine studies using the MMSE against the ACE or ACE-R generated a pooled sensitivity of 92.0% (95% CI = 84.9% to 96.8%) and specificity of 86.9% (95% CI = 80.5% to 92.1%) for the MMSE.

Comparison of ACE, ACE-R and MMSE at equivalent prevalence

Table 4 shows a direct comparison of the ACE, ACE-R, and MMSE at equivalent prevalence rates of 25% and 50%.

At 25% the overall accuracy of the ACE would be 0.823, the ACE-R 0.895, and the MMSE 0.882 (0.858 in the five studies compared with ACE alone). At 25% prevalence the PPV and NPV of the ACE would be 58.9% and 98.8% compared with 71.9% and 98.3% for the ACE-R, and 70.1% and 97.0% for the MMSE.

At 50% prevalence, the overall accuracy of the ACE would be 0.872, that of the ACE-R 0.916 and the MMSE 0.895 (0.838 in the four studies compared with ACE-R alone). The PPV and NPV of the ACE would be 81.1% and 96.1% compared with 88.5% and 95.3% for the ACE-R, and 87.5% and 91.6% for the MMSE.

According to the clinical utility index (see www.clinicalutility.co.uk), the ACE-R would have good utility at 25% prevalence and excellent properties at 50% prevalence, while the ACE would be qualitatively rated as good only at 50% prevalence. Figure 2 shows the accuracy at all possible prevalences and for rule-in (case-finding) and

Table 4. Comparison of ACE, ACE-R, and MMSE at difference prevalence (% with 95% confidence intervals)

	25% PREVALENCE			50% PREVALENCE		
	ACE	ACE-R	MMSE*	ACE	ACE-R	MMSE*
Overall accuracy**	0.823 (0.800–0.846)	0.895 (0.878–0.913)	0.882	0.872 (0.851–0.891)	0.916 (0.900–0.932)	0.895
Pooled Sensitivity	0.969 (0.954–0.983)	0.955 (0.931–0.979)	0.920 (0.903–0.937)	0.969 (0.954–0.983)	0.957 (0.940–0.973)	0.920 (0.903–0.937)
Pooled specificity	0.774 (0.745–0.802)	0.875 (0.853–0.897)	0.869 (0.857–0.881)	0.774 (0.739–0.809)	0.875 (0.849–0.902)	0.869 (0.848–0.890)
PPV	0.589 (0.543–0.634)	0.719 (0.674–0.764)	0.701 (0.676–0.725)	0.811 (0.781–0.841)	0.885 (0.860–0.910)	0.875 (0.855–0.895)
NPV	0.988 (0.979–0.996)	0.983 (0.974–0.992)	0.970 (0.964–0.977)	0.961 (0.943–0.979)	0.953 (0.935–0.971)	0.916 (0.898–0.933)
LR+	4.29 (3.77–4.87)	7.67 (6.42–9.16)	7.02 (6.39–7.71)	4.29 (3.67–5.02)	7.68 (6.18–9.54)	7.02 (5.98–8.25)
LR-	0.04 (0.02–0.08)	0.05 (0.03–0.09)	0.09 (0.07–0.11)	0.04 (0.03–0.06)	0.05 (0.03–0.07)	0.09 (0.07–0.11)
CUI+	0.572 (fair)	0.686 (good)	0.645 (good)	0.786 (good)	0.847 (excellent)	0.805 (good)
CUI-	0.764 (good)	0.861 (excellent)	0.843 (excellent)	0.744 (good)	0.834 (excellent)	0.796 (good)

* based on comparison data with ACE ($n = 5$)
** TP+TN/TP+TN+FP+FN

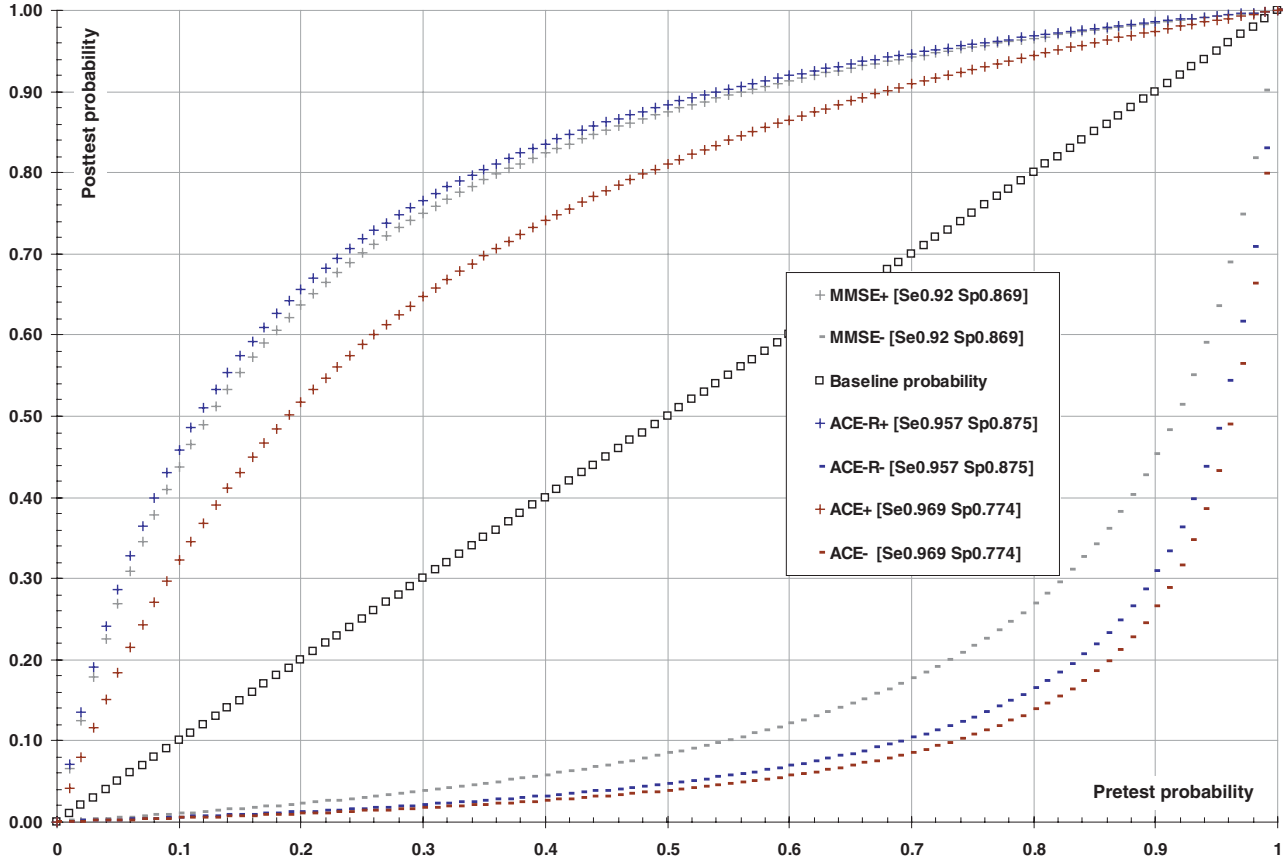

Figure 2. (Colour online) Conditional probability plot of PPV and NPV at fixed sensitivity and specificity.

rule-out (screening) purposes separately. Overall the ACE-R and MMSE are almost equivalent in terms of rule-in accuracy but the ACE is somewhat inferior. The ACE and ACE-R have approximately the same rule-out ability whilst the MMSE is somewhat inferior. When using one tool for both ruling-in cases and ruling-out non-cases (as is common practice), then only the ACE-R is recommended.

Discussion

In this meta-analysis, we examined the diagnostic accuracy of the ACE and ACE-R. Of 61 studies identified, only 10 (= 16.4%) fulfilled inclusion criteria; in a similar meta-analysis of MMSE diagnostic accuracy, of 775 studies identified, 34 (= 4.4%) fulfilled inclusion criteria (Mitchell, 2009). Studies included here varied in terms of subjects (i.e. non-demented groups variously encompassed healthy controls, those with subjective memory impairment, and those with cognitive impairment without dementia) and test cut-offs used.

Study limitations included the small number of studies suitable for meta-analysis and the possibility of bias in these studies. Since the ACE and ACE-R may presently be superseded by ACE-III (Hsieh *et al.*, 2013), it would seem unlikely that further large diagnostic accuracy studies of ACE and ACE-R suitable for inclusion in a meta-analysis will be forthcoming. As regards bias, use of normal control groups in some studies may have inflated test metrics. Although some studies clearly used non-overlapping patient cohorts (e.g. Larner 2007, 2013a), it is possible that the same participants were included in some studies (Yoshida *et al.* 2011, 2012). Data from different translations of the ACE/ACE-R used to test patients from different cultures may also be problematic. No studies were identified to permit analysis of ACE/ACE-R diagnostic accuracy for MCI.

From the ten valid studies using these tools (five each for ACE and ACE-R) to identify dementia compared with mixed subjects, both ACE and ACE-R were highly sensitive (> 95%), more so than simultaneous MMSE (< 90%). In terms of specificity, ACE-R was best (87.5%), followed by MMSE (81.7%) and ACE (77.4%). In a modest prevalence setting, such as primary care or general hospital settings, where the prevalence of dementia may be approximately 25%, the overall accuracy of the ACE would be lower than that of the MMSE

and the ACE-R. While both ACE and ACE-R would have adequate rule-out performance (see Figure 2), the PPV of the ACE would be less than that of the ACE-R (58.9% compared with 71.9%). Thus the ACE cannot be recommended in low prevalence settings, rather the ACE-R is preferred.

In high prevalence settings such as memory clinics, where dementia prevalence may be 50% or higher, both the ACE and ACE-R can be used but again the ACE-R is preferable. In everyday terms, of 20 above threshold scorers on the ACE-R in a high prevalence setting only one person would have dementia (false negative); of ten under threshold scorers, only one would not have dementia (false positive).

The results for the MMSE in specialist settings may be compared with those found in a larger meta-analysis of MMSE diagnostic accuracy studies for dementia ($n = 26$) which showed sensitivity and specificity of 76.9% (95% CI = 70.1% to 83.1%) and 89.9% (95% CI = 82.5% to 95.4%) (Mitchell, 2013). Results of the current study suggest much improved sensitivity for the MMSE (92.0% vs 76.9%), the interpretation of which is uncertain but might possibly reflect improved accuracy of the MMSE in recent studies, perhaps related to case selection in these studies. The close performance of the MMSE and ACE/ACE-R may not be surprising given that the ACE and ACE-R incorporate all of the MMSE stem questions (Table 1), unlike the ACE-III (Hsieh *et al*., 2013).

Results from a National Dementia Register suggested that ACE-R items improved MMSE estimates of cognitive ability by 16% (Law *et al*., 2013). Tests may be compared using the area under the curve of receiver operating characteristic (ROC) curves (AUC), but this method has been criticized in favor of weighted comparisons (Mallett *et al*., 2012). Such a weighted comparison of pragmatic studies suggested a net benefit for ACE-R versus MMSE and an equivalent increase in true positive patients of 121 per 1,000 tested (Larner, 2013b). Clearly the MMSE is briefer and quicker to administer than ACE or ACE-R; there is inevitably a trade-off between time and accuracy with cognitive screening instruments, and so MMSE might be suitable for use in time-limited situations, particularly as a rule-out, initial screen.

ACE-III has only recently been published (Hsieh *et al*., 2013) and is yet to be robustly validated. At the current time, the evidence of this meta-analysis favors the ACE-R over ACE and MMSE in the identification of dementia.

Conflict of interest

None.

Description of authors' roles

AJ Mitchell conceived and planned the study. Both authors abstracted the data, evaluated publications against the listed criteria, and resolved any disagreements about inclusion/exclusion through discussion. AJ Mitchell analyzed the data. AJ Larner drafted the manuscript. Both authors revised the manuscript for intellectual content.

References

Alexopoulos, P. *et al.* (2010). Validation of the German revised Addenbrooke's Cognitive Examination for detecting mild cognitive impairment, mild dementia in Alzheimer's disease and frontotemporal lobar degeneration. *Dementia and Geriatric Cognitive Disorders*, 29, 448–456. doi: 10.1159/000312685.

Crawford, S., Whitnall, L., Robertson, J. and Evans, J. J. (2012). A systematic review of the accuracy and clinical utility of the Addenbrooke's Cognitive Examination and the Addenbrooke's Cognitive Examination-Revised in the diagnosis of dementia. *International Journal of Geriatric Psychiatry*, 27, 659–669. doi: 10.1002/gps.2771.

Davies, R. R. and Larner, A. J. (2013). Addenbrooke's Cognitive Examination (ACE) and its Revision (ACE-R). In A. J. Larner (ed.), *Cognitive Screenisng Instruments. A Practical Approach* (pp. 61–77). London: Springer.

Dos Santos Kawata, K. H. *et al.* (2012). A validation study of the Japanese version of the Addenbrooke's Cognitive Examination-Revised. *Dementia and Geriatric Cognitive Disorders Extra*, 2, 29–37. doi: 10.1159/000336909.

Folstein, M. F., Folstein, S. E. and McHugh, P. R. (1975). "Mini-Mental State." A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189–198.

Garcia-Caballero, A. *et al.* (2006). Validation of the Spanish version of the Addenbrooke's Cognitive Examination in a rural community in Spain. *International Journal of Geriatric Psychiatry*, 21, 239–245.

Harbord, R. M., Egger, M. and Sterne, J. A. (2006). A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. *Statistics in Medicine*, 25, 3443–3457.

Hsieh, S., Schubert, S., Hoon, C., Mioshi, E. and Hodges, J. R. (2013). Validation of the Addenbrooke's Cognitive Examination-III in frontotemporal dementia and Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, 36, 242–250. doi: 10.1159/000351671.

Larner, A. J. (2007). Addenbrooke's Cognitive Examination (ACE) for the diagnosis and differential diagnosis of dementia. *Clinical Neurology and Neurosurgery*, 109, 491–494.

Larner, A. J. (2013a). Addenbrooke's Cognitive Examination-Revised (ACE-R): pragmatic study of cross-sectional use for assessment of cognitive complaints of unknown aetiology. *International Journal of Geriatric Psychiatry*, 28, 547–548. doi: 10.1002/gps.3884.

Larner, A. J. (2013b). Comparing diagnostic accuracy of cognitive screening instruments: a weighted comparison

approach. *Dementia and Geriatric Cognitive Disorders Extra*, 3, 60–65. doi: 10.1159/000348623.

Law, E. *et al.* (2013). Does the Addenbrooke's Cognitive Examination-Revised add to the Mini-Mental State Examination in established Alzheimer disease? Results from a national dementia research register. *International Journal of Geriatric Psychiatry*, 28, 351–355. doi: 10.1002/gps.3828.

Liberati, A. *et al.* (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ*, 339, b2700. doi: 10.1136/bmj.b2700.

Mallett, S., Halligan, S., Thompson, M., Collins, G. S. and Altman, D. G. (2012). Interpreting diagnostic accuracy studies for patient care. *BMJ*, 344, e3999. doi: 10.1136/bmj.e3999.

Mathuranath, P. S., Nestor, P. J., Berrios, G. E., Rakowicz, W. and Hodges, J. R. (2000). A brief cognitive test battery to differentiate Alzheimer's disease and frontotemporal dementia. *Neurology* 55, 1613–1620.

Mioshi, E., Dawson, K., Mitchell, J., Arnold, R. and Hodges, J. R. (2006). The Addenbrooke's Cognitive Examination Revised: a brief cognitive test battery for dementia screening. *International Journal of Geriatric Psychiatry*, 21, 1078–1085.

Mitchell, A. J. (2009). A meta-analysis of the accuracy of the Mini-Mental State Examination in the detection of dementia and mild cognitive impairment. *Journal of Psychiatric Research*, 43, 411–431. doi:10.1016/j.jpsychires.2008.04.014.

Mitchell, A. J. (2013). The Mini-Mental State Examination (MMSE): an update on its diagnostic validity for cognitive disorders. In A. J. Larner (ed.), *Cognitive Screening Instruments. A Practical Approach* (pp. 15–46). London: Springer.

Moher, D. *et al.* (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ*, 339, b2535. doi: 10.1136/bmj.b2535.

Stokholm, J., Vogel, A., Johannsen, P. and Waldemar, G. (2009). Validation of the Danish Addenbrooke's Cognitive Examination as a screening test in a memory clinic. *Dementia and Geriatric Cognitive Disorders*, 27, 361–365. doi: 10.1159/000209271.

Yoshida, H. *et al.* (2011). Validation of Addenbrooke's cognitive examination for detecting early dementia in a Japanese population. *Psychiatry Research*, 185, 211–214. doi: 10.1016/j.psychres.2009.06.012.

Yoshida, H. *et al.* (2012). Validation of the Revised Addenbrooke's Cognitive Examination (ACE-R) for detecting mild cognitive impairment and dementia in a Japanese population. *International Psychogeriatrics*, 24, 28–37. doi: 10.1017/S1041610211001190.