

This is an Open Access article licensed under the terms of the Creative Commons Attribution-NonCommercial 3.0 Unported license (CC BY-NC) (www.karger.com/OA-license), applicable to the online version of the article only. Distribution permitted for non-commercial purposes only.

Original Research Article

The Q* Index: A Useful Global Measure of Dementia Screening Test Accuracy?

A.J. Larner

Cognitive Function Clinic, Walton Centre for Neurology and Neurosurgery, Liverpool, UK

Key Words

Cutoff · Screening accuracy · Mini-Mental State Examination · Addenbrooke's Cognitive Examination-Revised · Montreal Cognitive Assessment · Test Your Memory test · Mini-Addenbrooke's Cognitive Examination · Q* index

Abstract

Background/Aims: Single, global or unitary, indicators of test diagnostic performance have intuitive appeal for clinicians. The Q* index, the point in receiver operating characteristic (ROC) curve space closest to the ideal top left-hand corner and where test sensitivity and specificity are equal, is one such measure. **Methods:** Datasets from four pragmatic accuracy studies which examined the Mini-Mental State Examination, Addenbrooke's Cognitive Examination-Revised, Montreal Cognitive Assessment, Test Your Memory test, and Mini-Addenbrooke's Cognitive Examination were examined to calculate and compare the Q* index, the maximal correct classification accuracy, and the maximal Youden index, as well as the sensitivity and specificity at these cutoffs. **Results:** Tests ranked similarly for the Q* index and the area under the ROC curve (AUC ROC). The Q* index cutoff was more sensitive (and less specific) than the maximal correct classification accuracy cutoff, and less sensitive (and more specific) than the maximal Youden index cutoff. **Conclusion:** The Q* index may be a useful global parameter summarising the test accuracy of cognitive screening instruments, facilitating comparison between tests, and defining a possible test cutoff value. As the point of equal sensitivity and specificity, its use may be more intuitive and appealing for clinicians than AUC ROC.

© 2015 S. Karger AG, Basel

Introduction

Diagnostic test accuracy is usually expressed in terms of test sensitivity and specificity, and sometimes also as positive and negative likelihood ratios. Other, single, global or unitary, indicators of test diagnostic performance have also been described, including overall test accuracy or correct classification accuracy; the Youden index (a combination of sensitivity

A.J. Larner
Cognitive Function Clinic
Walton Centre for Neurology and Neurosurgery
Lower Lane, Fazakerley, Liverpool L9 7LJ (UK)
E-Mail a.larner@thewaltoncentre.nhs.uk

and specificity, given by sensitivity + specificity – 1); the diagnostic odds ratio or cross-product ratio; the area under the receiver operating characteristic curve (AUC ROC), and measures of effect size such as Cohen's d [1]. Some of these global parameters (accuracy, Youden index, diagnostic odds ratio) are ultimately dependent upon the investigators' choice of test cutoff, cutpoint, threshold, or dichotomisation point, whereas other measures (AUC ROC and Cohen's d) are independent of this.

Another potentially useful summary measure denoting the diagnostic value of a test is the Q* index derived from the ROC curve [2]. The Q* index is defined as the 'point of indifference on the ROC curve', where the sensitivity and specificity are equal, or, in other words, where the probabilities of incorrect test results are equal for disease cases and non-cases (i.e. indifference between false-positive and false-negative diagnostic errors, with both assumed to be of equal value). The Q* index is that point in the ROC space which is closest to the ideal top left-hand corner of the ROC curve, where the anti-diagonal through the ROC space intersects the ROC curve [2].

An example of the potential utility of the Q* index was given by Hu et al. [3] who performed a meta-analysis of diagnostic accuracy studies of the International HIV Dementia Scale and the HIV Dementia Scale and found Q* index values of 0.9195 and 0.6321, respectively, for these tests. Although initially defined as a meta-analytic tool [2], the Q* index might also be applicable to individual diagnostic test accuracy studies for comparative purposes and to define the test cutoff point.

The primary aim of the study presented here was to examine the Q* index for a number of dementia cognitive screening instruments (CSIs) examined in pragmatic test accuracy studies and to compare the ranking of these CSIs according to either the AUC ROC or the Q* index. Admittedly, as their name implies, CSIs are not diagnostic tests for dementia since they do not address the underlying biology of disease; rather, they are aids to patient examination. Nevertheless, both screening and diagnostic tests require assessment using test accuracy studies, and these typically share the same methodology [1]. Moreover, since sophisticated diagnostic biomarker tests for dementia may not be universally, or indeed widely, available, the use of CSIs is likely to persist in clinical practice for some time. Hence, it would seem legitimate, and potentially clinically informative, to undertake comparative studies of CSIs examining parameters of test accuracy.

In addition, since Q* is a point, it may be used to define the test cutoff. Accordingly, the secondary aim of this study was to compare test sensitivity and specificity at the Q* index cutoff point with those values obtained using different test cutoffs, specifically those cited in index studies, and those derived from either maximal correct classification accuracy or the maximal Youden index, as previously examined [4]. The prediction was that for very sensitive tests, the Q* index would show lower sensitivity than that using other test cutoff methods, whereas for very specific tests, it would show better sensitivity.

Materials and Methods

The datasets from four pragmatic prospective test accuracy studies undertaken in dedicated secondary care memory clinics [5–8] were used, which examined five different CSIs: the Mini-Mental State Examination (MMSE) [9], the Addenbrooke's Cognitive Examination-Revised (ACE-R) [10], the Montreal Cognitive Assessment (MoCA) [11], the Test Your Memory (TYM) test [12], and the Mini-Addenbrooke's Cognitive Examination (M-ACE) [13]. Study details (setting, sample size, dementia prevalence, sex ratio, and age range) are shown in table 1. The Q* index was derived empirically from ROC curves based on study data (fig. 1), as were the maximal correct classification accuracy and the maximal Youden index [4].

Table 1. Study demographics

CSI	Setting	n	Dementia prevalence	F:M (% females)	Age range, years	Ref.
MMSE	cognitive function clinic	242	0.35	108:134 (45)	24–85 (mean 60)	5
ACE-R	cognitive function clinic	243	0.35	108:135 (44)	24–85 (mean 60)	5
MoCA	cognitive function clinic	150	0.24	57:93 (38)	20–87 (median 61)	6
TYM	cognitive function clinic and old age psychiatry memory clinic	224	0.35	94:130 (42)	20–90 (mean 63)	7
M-ACE	cognitive function clinic	135	0.18	64:71(47)	18–88 (median 60)	8

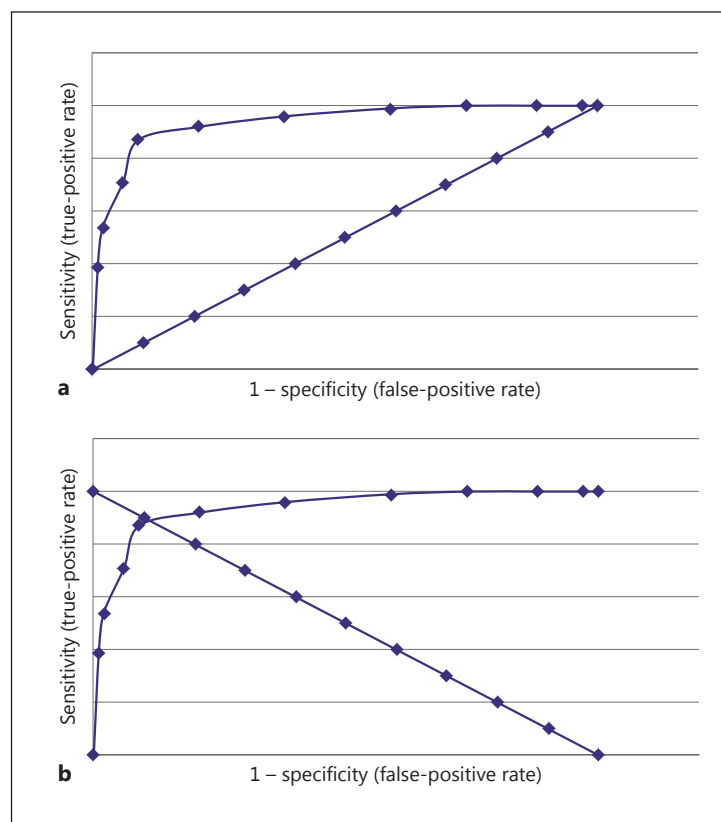


Fig. 1. a Typical ROC curve or plot with diagonal or chance line (data for ACE-R adapted from [5]). **b** Typical ROC curve (same data points as in **a**) with anti-diagonal line: where the lines cross in the ROC space indicates equal test sensitivity and specificity, by definition the Q* index (the point closest to the ideal top left-hand corner of the ROC curve).

Results

The Q* index ranged from 0.88 for ACE-R to 0.76 for M-ACE (table 2). The ranking of the Q* index for the various CSIs examined paralleled that for the AUC ROC, with ACE-R ranking highest and M-ACE lowest using either parameter.

Comparing the Q* index cutoff point with cutoffs defined in index studies, the former was always lower (and hence less sensitive but more specific) than the latter (table 3).

Comparing test sensitivity and specificity at the Q* index cutoff point showed that for all CSIs with the exception of ACE-R, Q* index-derived test cutoffs lay between those derived from the maximal correct classification accuracy and the maximal Youden index (table 3).

Table 2 Summary 'league table' of the AUC ROC and the Q* index for various CSIs

Test	AUC ROC (95% CI) (ranking)	Q* index (ranking)
MMSE	0.91 (0.88–0.95) (2=)	0.82 (2)
ACE-R	0.94 (0.91–0.97) (1)	0.88 (1)
MoCA	0.91 (0.86–0.95) (2=)	0.79 (4)
TYM	0.89 (0.84–0.93) (4)	0.80 (3)
M-ACE	0.86 (0.83–0.90) (5)	0.76 (5)

Table 3. Comparison of the sensitivity and specificity for different CSIs examined in pragmatic diagnostic test accuracy studies using cutoffs adopted from index studies and cutoffs defined by the Q* index, the maximal correct classification accuracy, and the maximal Youden index

	Cutoff	Sensitivity	Specificity
MMSE			
Q* index	≥25/30	0.82	0.82
Maximal accuracy	≥24/30	0.70	0.89
Maximal Youden index	≥26/30	0.92	0.75
ACE-R			
Index study [10]	≥88/100	0.99	0.44
Index study [10]	≥82/100	0.96	0.62
Q* index	≥74/100	0.88	0.88
Maximal accuracy = maximal Youden index	≥73/100	0.87	0.91
MoCA			
Index study [11]	≥26/30	0.97	0.60
Q* index	≥23/30	0.79	0.79
Maximal accuracy	≥20/30	0.63	0.95
Maximal Youden index	≥24/30	0.86	0.76
TYM			
Index study [12]	≤42/50	0.95	0.45
Q* index	≤34/50	0.80	0.80
Maximal accuracy	≤30/50	0.73	0.88
Maximal Youden index	≤37/50	0.91	0.72
M-ACE			
Index study [13]	≤25/30	1.00	0.28
Index study [13]	≤21/30	0.92	0.61
Q* index	≤17/30	0.76	0.76
Maximal accuracy	≤13/30	0.46	0.93
Maximal Youden index	≤20/30	0.92	0.69

Hence, if the Q* index point were used as the test cutoff, it was more sensitive (and less specific) than if using the maximal correct classification accuracy cutoff, and less sensitive (and more specific) than if using the maximal Youden index cutoff.

Q* index cutoffs reduced the sensitivity of very sensitive tests such as the ACE-R, MoCA, TYM, and M-ACE ≤25/30, but improved the sensitivity for very specific tests such as M-ACE ≤21/30.

Discussion

Q* index values were smaller than AUC ROC values, as previously observed [2, 3], but showed a greater range than AUC ROC. This range might relate in part to the different coverage of cognitive domains in each of the tests, a factor which may impact on test accuracy [14].

The various possible cutoffs each have their own advantages and disadvantages for diagnostic purposes. The Q* index assumes that false-positive and false-negative diagnostic errors are of equal value, as does the maximal correct classification accuracy, an assumption which may be moot (clinicians may be unwilling to countenance false negatives or missed diagnoses). Correct classification accuracy is dependent on disease prevalence, unlike the Youden index, although the latter does arbitrarily assume the disease prevalence to be 50%, which is seldom likely to be the case. The Q* index is the point closest to the ideal top left-hand corner of the ROC curve, whilst the maximal Youden index corresponds to the maximal vertical distance between the ROC curve and the diagonal or chance line (i.e. the point on the curve furthest from chance).

One possible shortcoming of the current analysis is the comparison of CSIs across different studies, since these were not (and never could be) exactly comparable. Nevertheless, the studies being compared here shared a similar setting, methodology (e.g. application of reference standard for diagnosis), and analysis, thereby minimising variability, although this cannot be entirely excluded (e.g. dementia prevalence, based on the casemix seen in series of unselected consecutive patients). Moreover, it is certainly the case that different CSIs may, according to clinician judgement, be appropriate in different clinical scenarios, irrespective of the outcomes of comparative studies. Nevertheless, such comparative studies may demonstrate the potential advantages and disadvantages of particular CSIs but without undermining clinicians' ultimate autonomy in choice of test.

Whether or not the Q* index is adopted will depend on a number of factors. Currently, the Q* index remains an unfamiliar metric, and clinicians may be unwilling to countenance its use in place of more familiar metrics such as sensitivity, specificity, and predictive values, even though all of these also have their pros and cons [1]. A broadening of clinician literacy would likely be required for the widespread adoption of the Q* index as a measure of test discrimination. It is not argued that the Q* index is better than existing methods, but that it may add information to that available from other parameters and inform clinicians' choice of screening test.

Another issue determining or influencing test choice relates to what precisely investigators want from the tests they administer. If investigators want to maximise test sensitivity (fewest false negatives or missed diagnoses), then a cutoff based on the maximal Youden index may be more appropriate than one based on the Q* index, whereas if they want to maximise specificity (fewest false positives), then a cutoff based on maximal correct classification accuracy might be chosen. However, if false-positive and false-negative diagnoses are deemed equally undesirable, the Q* index cutoff should be chosen.

If a metric to compare diagnostic tests is required, the Q* index has merit and, since it is based on sensitivity and specificity, may perhaps be preferred to AUC ROC results as more intuitive. It has been argued that ROC curves 'have little intuitive appeal for physicians' [15], and, moreover, AUC ROC may be criticised for combining test accuracy over a range of thresholds which may be both clinically relevant and clinically nonsensical [16].

Disclosure Statement

The author declares no conflicts of interest.

References

- 1 Larner AJ: Diagnostic Test Accuracy Studies in Dementia. A Pragmatic Approach. London, Springer, 2015.
- 2 Walter SD: Properties of the summary receiver operating characteristic (SROC) curve for diagnostic test data. *Stat Med* 2002;21:1237–1256.
- 3 Hu X, Zhou Y, Long J, Feng Q, Wang R, Su L, Zhao T, Wei B: Diagnostic accuracy of the International HIV Dementia Scale and HIV Dementia Scale: a meta-analysis. *Exp Ther Med* 2012;4:665–668.
- 4 Larner AJ: Optimizing the cutoffs of cognitive screening instruments in pragmatic diagnostic accuracy studies: maximising accuracy or Youden index? *Dement Geriatr Cogn Disord* 2015;39:167–175.
- 5 Larner AJ: Addenbrooke's Cognitive Examination-Revised (ACE-R): pragmatic study of cross-sectional use for assessment of cognitive complaints of unknown aetiology. *Int J Geriatr Psychiatry* 2013;28:547–548.
- 6 Larner AJ: Screening utility of the Montreal Cognitive Assessment (MoCA): in place of – or as well as – the MMSE? *Int Psychogeriatr* 2012;24:391–396.
- 7 Hancock P, Larner AJ: Test Your Memory (TYM) test: diagnostic utility in a memory clinic population. *Int J Geriatr Psychiatry* 2011;26:976–980.
- 8 Larner AJ: Mini-Addenbrooke's Cognitive Examination: a pragmatic diagnostic accuracy study. *Int J Geriatr Psychiatry* 2015;30:547–548.
- 9 Folstein MF, Folstein SE, McHugh PR: 'Mini-Mental State'. A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12:189–198.
- 10 Mioshi E, Dawson K, Mitchell J, Arnold R, Hodges JR: The Addenbrooke's Cognitive Examination Revised: a brief cognitive test battery for dementia screening. *Int J Geriatr Psychiatry* 2006;21:1078–1085.
- 11 Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, Cummings JL, Chertkow H: The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc* 2005;53:695–699.
- 12 Brown J, Pengas G, Dawson K, Brown LA, Clatworthy P: Self-administered cognitive screening test (TYM) for detection of Alzheimer's disease: cross-sectional study. *BMJ* 2009;338:b2030.
- 13 Hsieh S, McGrory S, Leslie F, Dawson K, Ahmed S, Butler CR, Rowe JB, Mioshi E, Hodges JR: The Mini-Addenbrooke's Cognitive Examination: a new assessment tool for dementia. *Dement Geriatr Cogn Disord* 2015;39:1–11.
- 14 Larner AJ: Speed versus accuracy in cognitive assessment when using CSIs. *Prog Neurol Psychiatry* 2015;19(1):21–24.
- 15 Richard E, Schmand BA, Eikelenboom P, Van Gool WA, The Alzheimer's Disease Neuroimaging Initiative: MRI and cerebrospinal fluid biomarkers for predicting progression to Alzheimer's disease in patients with mild cognitive impairment: a diagnostic accuracy study. *BMJ Open* 2013;3:e002541.
- 16 Mallett S, Halligan S, Thompson M, Collins GS, Altman DG: Interpreting diagnostic accuracy studies for patient care. *BMJ* 2012;344:e3999.