Evaluating binary classifiers: extending the Efficiency Index

Andrew J Larner Consultant Neurologist

Cognitive Function Clinic, Walton Centre for Neurology and Neurosurgery,
Liverpool, United Kingdom

Correspondence: AJ Larner, Cognitive Function Clinic, Walton Centre for Neurology
and Neurosurgery, Lower Lane, Fazakerley, Liverpool, L9 7LJ, United Kingdom
e-mail: a.larner@thewaltoncentre.nhs.uk

Abstract

An “efficiency index” (EI) for the evaluation of binary classifiers was recently characterised, where EI is the ratio of classifier accuracy to inaccuracy. The purpose of this study was to further develop EI by substituting balanced accuracy and unbiased accuracy in place of accuracy, and their respective complements in place of inaccuracy, to construct balanced EI and unbiased EI measures. Additional investigations, using the dataset of a prospective pragmatic test accuracy study of a cognitive screening instrument, explored use of the log method to calculate confidence intervals for the various EI formulations; the dependence of EI formulations on prevalence; and comparison of EI formulations with analogous formulations based on the Identification Index (II), a previously described metric which is also based on accuracy and inaccuracy, where II is accuracy minus inaccuracy. EI formulations are shown to have advantages over II formulations, in particular their boundary values (0 and ∞) mean that negative values never occur, unlike the case for II, and the inflection point of 1 demarcates likelihood of correct versus incorrect classification.

Keywords: binary classifier; efficiency indexes; identification indexes; screening test

1. Introduction

Many metrics exist for the evaluation of binary classifiers, with different measures being favoured in different disciplines. For example, the true positive and true negative rates (TPR, TNR) derived from the standard 2x2 contingency table, also known as Sensitivity and Specificity, are often used in the evaluation of medical screening and diagnostic tests,¹ whereas the F measure, also known as F or F1 score,² and Matthews' correlation coefficient (MCC)³ are often preferred in data retrieval and machine learning contexts. Unlike the paired measures of TPR and TNR, F measure and MCC are unitary, global measures of test accuracy, other examples of which include Accuracy (Acc) and the area under the receiver operating characteristic curve (AUC ROC).

Recently a further metric, the efficiency index (EI), was proposed as a measure to evaluate binary classifiers.⁴ (Note that this EI differs from the similarly named physical index of speed achieved in relation to power output, and various other efficiency indexes described for energy efficiency or in business and finance.) EI was based on Acc and its complement, Inaccuracy (Inacc), and defined as:

$$EI = \frac{Acc}{Inacc}$$

Acc is a weighted average of TPR and TNR, with weights related to sample prevalence (P and 1 - P or P'):

$$Acc = TPR.P + TNR.P'$$

Similarly, Inacc is a weighted average of false positive rate (FPR) and false negative rate (FNR), with weights related to P and P':

$$Inacc = FNR.P + FPR.P'$$

where FNR and FPR are the complements of TPR and TNR respectively. Hence, EI may be rewritten as

$$EI = \frac{TPR.P + TNR.P'}{FNR.P + FPR.P'}$$

which indicates the dependence of EI on prevalence.

EI has been suggested to have some advantages over other unitary measures of binary classification.⁴ EI has boundary values of 0 and ∞ , with an inflection point at 1, where a value >1 indicates correct classification and a value of <1 indicates incorrect classification, such that values >>1 are desirable and a value of ∞ is optimal. These boundary values and inflection point are shared with likelihood ratios (LRs) and it has previously been shown that EI may be characterised using the same qualitative and quantitative systems of classification used for LRs.^{5,6} This may be of advantage when communicating risk, for example of a medical test producing correct diagnosis rather than misdiagnosis.⁴

The purpose of this study was to extend consideration of EI in various ways. Firstly, by using other formulations of Acc, namely Balanced Accuracy and Unbiased Accuracy,⁷ new Balanced EI and Unbiased EI formulations have been developed.

Secondly, confidence intervals for these EI formulations have been calculated using the same logarithmic method used for LR's and for diagnostic odds ratios (DORs). Thirdly, the dependence of EI formulations on prevalence has been examined. Fourthly, the EI metrics have been compared with another, previously described, metric based on Acc and Inacc, the Identification Index (II),⁸ which was also further developed using Balanced Accuracy and Unbiased Accuracy to produce Balanced II and Unbiased II formulations. Worked examples of all these developments of EI are examined using the dataset of a test accuracy study for a dementia screening test.

2. Material and methods

The utility of EI was previously shown⁴ using published data from a prospective test accuracy study of a screening test for dementia and cognitive impairment, the Mini-Addenbrooke's Cognitive Examination (MACE).⁹ The same dataset¹⁰ was used here, from a study in which subjects gave informed consent and study protocol was approved by the institute's committee on human research (Walton Centre for Neurology and Neurosurgery Approval: N 310).

2.1 Extended formulations of EI

Balanced Accuracy (BAcc) has been defined¹¹ as:

$$\text{BAcc} = (\text{TPR} + \text{TNR})/2$$

BAcc has been advocated in place of Acc when examining biased classifiers and/or imbalanced datasets, i.e. where P and hence P' are markedly different from 0.5.^{7,11}

Using BAcc in place of Acc, one may characterise a Balanced Efficiency Index (BEI) as:

$$\begin{aligned} \text{BEI} &= \text{BAcc}/\text{BInacc} \\ &= \text{BAcc}/(1 - \text{BAcc}) \\ &= (\text{TPR} + \text{TNR})/(\text{FNR} + \text{FPR}) \end{aligned}$$

Hence unlike EI, BEI is independent of prevalence.

Unbiased Accuracy (UAcc) has been defined¹² as:

$$\text{UAcc} = (\text{TPR} \cdot \text{P} + \text{TNR} \cdot \text{P}') - (\text{P} \cdot \text{Q} + \text{P}' \cdot \text{Q}')/1 - (\text{P} \cdot \text{Q} + \text{P}' \cdot \text{Q}')$$

where Q is the level or bias of the test.¹³ This formulation corrects for the proportion of agreement expected by chance alone, equating to Cohen's kappa statistic, κ .^{7,12}

Using UAcc in place of Acc, one may characterise an Unbiased Efficiency Index (UEI) as:

$$\begin{aligned} \text{UEI} &= \text{UAcc}/\text{UInacc} \\ &= \text{UAcc}/(1 - \text{UAcc}) \end{aligned}$$

Like EI, BEI and UEI have boundary values of 0 and ∞ , with an inflection point at 1, where a value >1 indicates correct classification and a value of <1 indicates incorrect

classification, such that values $\gg 1$ are desirable and a value of ∞ is an optimal classifier.

2.2 Confidence intervals for EI formulations

EI, BEI, and UEI are all of the form $a/(1-a)$ and hence are odds ratios. (They could all also be expressed as $(1-a)/a$, hence as odds against ratios.) It is possible to calculate confidence intervals (CI) for EI, BEI, and UEI, as for LR and DORs, through a logarithmic transformation.¹⁴ Applying the log method to data from the four cells of the 2x2 contingency table (where true positive = a, false positive = b, false negative = c, true negative = d), for the 95% CI the formula for EI is:

$$\text{Log}_e(\text{EI}) \pm [1.96 \times \text{SE}(\log_e \text{EI})]$$

where:

$$\text{SE}(\log_e \text{EI}) = \sqrt{[1/a - 1/(a+c) + 1/b - 1/(b+d)]}$$

With the appropriate substitutions, this method may also be used to calculate 95% CI for BEI and UEI.

2.3 Dependence of EI formulations on prevalence

As BEI is based solely on columnar ratios from the 2x2 contingency table (TPR, FNR, TNR, FPR), it is independent of P. Whilst this might be deemed an advantage, the biasing effects of random associations between classifier result and prevalence may be allowed for by taking into account P, as is the case in EI and UEI. This may be examined by rescaling Acc and UAcc (and hence Inacc and UInacc) across a range of prevalence (P) values at a fixed value of Q (i.e. fixed classifier cut-off, and hence fixed values of TPR, TNR, FNR and FPR) and hence calculating rescaled EI and UEI values from the respective equations.

2.4 Comparison of EI formulations with II formulations

An Identification Index (II) has been defined⁸ as:

$$\begin{aligned} \text{II} &= \text{Acc} - \text{Inacc} \\ &= \text{Acc} - (1 - \text{Acc}) \\ &= 2 \cdot \text{Acc} - 1 \end{aligned}$$

Hence whereas Inacc is the divisor of Acc in EI, it is the subtrahend of Acc in II.

By analogy with EI and BEI, one may characterise a Balanced Identification Index (BII) as:

$$\begin{aligned} \text{BII} &= \text{BAcc} - \text{BInacc} \\ &= \text{BAcc} - (1 - \text{BAcc}) \\ &= 2 \cdot \text{BAcc} - 1 \end{aligned}$$

However, this formulation contributes no additional information, since:

$$\text{BAcc} = (\text{TPR} + \text{TNR})/2$$

Hence:

$$\begin{aligned} \text{BII} &= 2.[(\text{TPR} + \text{TNR})/2] - 1 \\ &= \text{TPR} + \text{TNR} - 1 \end{aligned}$$

This is the formula for the Youden index or J statistic,¹⁵ also termed bookmarker informedness,³ the value of which is immediately available from TPR and TNR.

By analogy with EI and UEI, one may also characterise an Unbiased Identification Index (UII) as:

$$\begin{aligned} \text{UII} &= \text{UAcc} - \text{UIInacc} \\ &= \text{UAcc} - (1 - \text{UAcc}) \\ &= 2.\text{UAcc} - 1 \end{aligned}$$

II, BII and UII have boundary values of -1 and +1, unlike the various efficiency indexes. They have an inflection point at 0, values above which imply the classifier is useful rather than useless, with a value of 0 denoting a random classifier, a value of 1 denoting an optimal classifier, and values <0 indicating that the classifier is misleading.

3. Results

3.1 Extended formulations of EI

Calculations of EI, BEI, and UEI are shown in Table 1 and illustrated in Figure 1. These parameters showed different maxima (EI at MACE cut-off $\leq 14/30$; BEI $\leq 20/30$; UEI $\leq 15/30$). The plots of EI and BEI cross (EI = BEI somewhere between the cut-offs of $\leq 18/30$ and $\leq 19/30$ at this prevalence) but neither cross the plot of UEI which lies beneath them. So stringent is UEI that it has a value >1 only at its maximum ($\leq 15/30$).

3.2 Confidence intervals for EI formulations

The 95% confidence intervals (95% CI) for EI, BEI, and UEI at the previously defined optimal MACE cut-off in the test dataset ($\leq 20/30$)¹⁰ were calculated by the log method as follows:

$$\begin{aligned} \text{EI} &= 2.82 \text{ (95\% CI = 2.46-3.21)} \\ \text{BEI} &= 4.24 \text{ (95\% CI = 3.71-4.84)} \\ \text{UEI} &= 0.61 \text{ (95\% CI = 0.53-0.69).} \end{aligned}$$

At their respective maxima, the values for EI and UEI were 6.55 (95% CI = 4.85-8.84) and 1.01 (95% CI = 0.78-1.32). Thus, even at its optimal cut-off, the 95% CI for UEI crossed 1.

3.3 Dependence of EI formulations on prevalence

The dependence of EI, BEI, and UEI on prevalence are shown in Table 2 and illustrated in Figure 2. By definition BEI is independent of P, hence does not vary, whereas EI and UEI gradually increase with increasing prevalence, since when $P = 1$ then EI and UEI = ∞ because the denominators in their respective defining equations

= 0, and when $P = 0$ then EI and UEI = 0 because the numerators in their respective defining equations = 0.

3.4 Comparison of EI formulations with II formulations

Calculations of II, BII, and UII are shown in Table 3 and illustrated in Figure 3 (note the different y-axis scale from Figure 1). These parameters showed different maxima ($II \leq 14/30$; $BII \leq 20/30$; $UII \leq 15/30$) which corresponded to the maxima for the corresponding EI formulations. The plots of II and BII cross ($II = BII = Y$ somewhere between the cut-offs of $\leq 18/30$ and $\leq 19/30$ at this prevalence) but neither cross the plot of UII which lies beneath them, as was observed for the various EI formulations. So stringent is UII that it has a value >0 only at its maximum ($\leq 15/30$).

4. Discussion

This study aimed to develop the recently proposed efficiency index, EI, as a metric for the evaluation of binary classifiers by application of different formulations of accuracy, namely balanced accuracy and unbiased accuracy. This has extended the formulation of EI to BEI and UEI by using BAcc and UAcc in place of Acc respectively. By the same procedure, the formulation of the identification index, II, has been extended to BII and UII.

Of the EI formulations, BEI was found to be a relatively more liberal or optimistic metric than EI, whereas UEI was a relatively more conservative or pessimistic metric, as was previously shown to be the case for UAcc relative to Acc.⁷ In the calculated example, only one UEI point had a value >1 , and this may be a more generalised problem because of the conservative values of UAcc. This arises because UEI, like UAcc, takes into account values of both P and Q, hence removing the biasing effects of random associations between, in this particular case, the test result and disease prevalence.

Confidence intervals for EI formulations were easily calculated from the 2x2 contingency table base data using the logarithmic method.¹⁴ In the worked examples, the confidence intervals for EI and BEI at the optimal MACE cut-off were greater than 1, suggesting that the test under examination was more likely to classify correctly than incorrectly, whereas for UEI even at its maximal cut-off the confidence intervals crossed 1, suggesting that when using this stringent metric this classifier is no more likely to classify correctly than to misclassify.

By definition BEI, like BAcc, is an unscaled measure, calculation of which is independent of P. EI and UEI formulations vary with P, by definition. This is because EI and UEI, like Acc and UAcc, take into account the value of P, hence removing the biasing effects of random associations between test result and disease prevalence. EI occupies a middle ground between BEI and UEI, since it takes into account P but not Q.⁴ In the worked examples, both EI and UEI increased with P for the classifier under examination. Using the same dataset, a previously described “likelihood to be diagnosed or misdiagnosed” (LDM), which has some overlap with EI,⁴ showed a fall in value at either very low or very high disease prevalence using the same cut-off.¹⁶

Of the II formulations, BII simplifies to an existing metric, the Youden index, Y , or bookmarker informedness, and hence contributes no new information. As it is based on strictly columnar ratios from the 2x2 contingency table (TPR, TNR), it is, at least algebraically, independent of P , although in practice TPR and TNR may be found to vary according to the heterogeneity of the sampled population.¹⁷ As UII is based on UAcc, it is by definition dependent on P and Q , with II occupying a middle ground (taking into account P but not Q).

EI formulations have advantages over II formulations. The range of the EI formulations (boundary values 0 and ∞) means that negative values never occur, as is also the case for LR and DORs. EI values are easily classifiable, either qualitatively or quantitatively,⁴ using existing classification schemes for LR,^{5,6} with an inflection point at 1, denoting a random classifier, above which correct classification is favoured over incorrect classification. The qualitative classification scheme for odds ratios proposed by Rosenthal might also be applicable to EI formulations (small $\square 1.5$, medium $\square 2.5$, large $\square 4$, and very large $\square 10$ or greater).¹⁸

In contrast to EI formulations, the range of the II formulations (boundary values -1 and +1) means that negative values can occur (see Table 3 and Figure 3), the meaning of which may be unclear. There is no evident classificatory scheme for these metrics, other than larger better, with an inflection point at 0, values above which imply the result is useful rather than useless, and below which imply the result is misleading. In the worked example, only one UII point had a value >0 , and this may be a more generalised problem because of the conservative values of UAcc. These shortcomings may account, at least in part, for the fact that the original II formulation has been rarely adopted or examined.^{1,15,19}

It has previously been suggested that EI may have greater utility than other metrics in communicating risk of test outcomes⁴ and this may also be true of the extended EI formulations developed here, although this requires empirical examination. Other future studies might also examine the relationship between EI formulations as possible indexes of effect size which might be qualitatively related to Cohen's d .²⁰

In summary, EI, BEI and UEI appear to be serviceable metrics for the evaluation of binary classifiers. They are easily calculated from the cells of the 2x2 contingency table, as are their confidence intervals using the log method. They are easily classified qualitatively and quantitatively. Of the various extensions of EI examined here, UEI is the most stringent in that it corrects for prevalence and threshold and therefore may prove to be the most appropriate of these measures to be used in the assessment and comparison of the performance of different classifiers.

Funding and Acknowledgements

None

References

1. Larner AJ. *The 2x2 matrix: contingency, confusion and the metrics of binary classification*. London: Springer; doi: [10.1007/978-3-030-74919-4](https://doi.org/10.1007/978-3-030-74919-4) (2021).
2. Powers DMW. What the F measure doesn't measure ... Features, flaws, fallacies and fixes. arXiv, doi: [1503.06410.2015](https://arxiv.org/abs/1503.06410) (2015).
3. Chicco D, Tötsch, N, Jurman G. The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markendness in two-class confusion matrix evaluation. *BioData Mining* **14**, 13, doi: [10.1186/s13040-021-00244-z](https://doi.org/10.1186/s13040-021-00244-z) (2021).
4. Larner AJ. Communicating risk: developing an "Efficiency Index" for dementia screening tests. *Brain Sci* **11**, 1473, doi: [10.3390/brainsci11111473](https://doi.org/10.3390/brainsci11111473) (2021).
5. Jaeschke R, Guyatt G, Sackett DL. Users' guide to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? *JAMA* **271**, 703-707, doi: [10.1001/jama.271.9.703](https://doi.org/10.1001/jama.271.9.703) (1994).
6. McGee S. Simplifying likelihood ratios. *J Gen Intern Med* **17**, 647-650, doi: [10.1046/j.1525-1497.2002.10750.x](https://doi.org/10.1046/j.1525-1497.2002.10750.x) (2002).
7. Larner AJ. Accuracy of cognitive screening instruments reconsidered: overall, balanced, or unbiased accuracy? *Neurodegener Dis Manag* **12**, accepted for publication (2022).
8. Mitchell AJ. Index test. In: Kattan MW, ed. *Encyclopedia of medical decision making*. Los Angeles, Sage; 613-617, doi: [10.4135/9781412971980](https://doi.org/10.4135/9781412971980) (2009).
9. Hsieh S, McGrory S, Leslie F, et al. The Mini-Addenbrooke's Cognitive Examination: a new assessment tool for dementia. *Dement Geriatr Cogn Disord* **39**, 1-11, doi: [10.1159/000366040](https://doi.org/10.1159/000366040) (2015).
10. Larner AJ. MACE for diagnosis of dementia and MCI: examining cut-offs and predictive values. *Diagnostics (Basel)* **9**, E51, doi: [10.3390/diagnostics9020051](https://doi.org/10.3390/diagnostics9020051) (2019).
11. Brodersen KH, Ong CS, Stephan KE, Buhmann JM. The balanced accuracy and its posterior distribution. In: Oliveira DM, Lins RD, eds. *Proceedings of ICPR 2010 – the 20th IAPR International Conference on Pattern Recognition*. IEEE; 3121-3124, doi: [10.1109/icpr.2010.590](https://doi.org/10.1109/icpr.2010.590) (2010).
12. Garrett CT, Sell S. Summary and perspective: Assessing test effectiveness – the identification of good tumour markers. In: Garrett CT, Sell S, eds. *Cellular cancer markers*. Totowa, NJ: Humana Press; 455-477, doi: [10.1007/978-1-4757-2381-6_17](https://doi.org/10.1007/978-1-4757-2381-6_17) (1995).
13. Kraemer HC. *Evaluating medical tests. Objective and quantitative guidelines*. Newbery Park, California: Sage (1992).
14. Katz D, Baptista J, Azen SP, Pike MC. Obtaining confidence intervals for the risk ratio in cohort studies. *Biometrics* **34**, 469-474, doi: [10.2307/2530610](https://doi.org/10.2307/2530610). (1978).
15. Youden WJ. Index for rating diagnostic tests. *Cancer* **3**, 32-35, doi: [10.1002/1097-0142\(1950\)3:1<32::AID-CNCR2820030106>3.0.CO;2-3](https://doi.org/10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3) (1950).
16. Larner AJ. *Manual of screeners for dementia. Pragmatic test accuracy studies*. London: Springer; doi: [10.1007/978-3-030-41636-2](https://doi.org/10.1007/978-3-030-41636-2) (2020).

17. Brenner H, Gefeller O. Variation of sensitivity, specificity, likelihood ratios and predictive values with disease prevalence. *Stat Med* **16**, 981-991, doi: [10.1002/\(sici\)1097-0258\(19970515\)16:9<981::aid-sim510>3.0.co;2-n](https://doi.org/10.1002/(sici)1097-0258(19970515)16:9<981::aid-sim510>3.0.co;2-n) (1997).
18. Rosenthal JA. Qualitative descriptors of strength of association and effect size. *J Soc Serv Res* **21**, 37-59, doi: [10.1300/j079v21n04_02](https://doi.org/10.1300/j079v21n04_02) (1996).
19. Mendez MF. *The mental status examination handbook*. Philadelphia: Elsevier; 88 (2022).
20. Chen H, Cohen P, Chen S. How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. *Commun Stat Simul Comput* **39**, 860-864, doi: [10.1080/03610911003650383](https://doi.org/10.1080/03610911003650383) (2010).

Table 1: Diagnosis of dementia: comparing EI, BEI, and UEI metrics at various MACE cut-offs (fixed value of P = 0.151)

Cut-off	EI	BEI	UEI
≤29/30	0.204	1.045	0.006
≤28/30	0.246	1.101	0.015
≤27/30	0.355	1.283	0.043
≤26/30	0.507	1.538	0.081
≤25/30	0.716	1.882	0.135
≤24/30	0.982	2.289	0.198
≤23/30	1.274	2.759	0.272
≤22/30	1.668	3.310	0.368
≤21/30	2.199	3.854	0.484
≤20/30	2.813	4.236	0.605
≤19/30	3.364	4.181	0.689
≤18/30	4.033	4.000	0.776
≤17/30	4.207	3.525	0.745
≤16/30	5.292	3.785	0.934
≤15/30	6.123	3.484	1.012
≤14/30	6.550	2.922	0.980
≤13/30	6.475	2.425	0.795
≤12/30	6.475	2.195	0.718
≤11/30	6.260	1.874	0.567

Abbreviations: EI = efficiency index, BEI = balanced efficiency index, UEI = unbiased efficiency index

Figure 1: Plot of EI (◆), BEI (▲) and UEI (×) values (y axis) vs MACE cut-off score (x-axis)

Table 2: Rescaled values of EI and UEI at varying values of P for diagnosis of dementia using MACE at fixed cut-off ($\leq 20/30$; $Q = 0.387$, $Q' = 0.613$) and hence Sens (0.912) and Spec (0.707), and hence fixed BEI (= 4.236)

P, P'	Rescaled Acc	Rescaled Inacc	Rescaled EI	Rescaled UAcc	Rescaled UInacc	Rescaled UEI
0.9, 0.1	0.892	0.108	8.216	0.816	0.184	4.435
0.8, 0.2	0.871	0.129	6.752	0.773	0.227	3.403
0.7, 0.3	0.851	0.149	5.689	0.726	0.274	2.645
0.6, 0.4	0.830	0.170	4.882	0.675	0.325	2.076
0.5, 0.5	0.809	0.191	4.249	0.619	0.381	1.625
0.4, 0.6	0.789	0.211	3.739	0.558	0.442	1.261
0.3, 0.7	0.769	0.231	3.319	0.491	0.509	0.965
0.2, 0.8	0.748	0.252	2.968	0.417	0.583	0.714
[0.151, 0.849]	0.738	0.262	2.813	0.377	0.623	0.605
0.1, 0.9	0.728	0.272	2.670	0.334	0.666	0.501

Abbreviations: EI = efficiency index, BEI = balanced efficiency index, UEI = unbiased efficiency index

Figure 2: Plot of rescaled EI (◆) and UEI (×) values with fixed BEI (▲) (y axis) vs prevalence (x-axis)

Table 3: Diagnosis of dementia: comparing II, BII, and UII metrics at various MACE cut-offs (fixed value of P = 0.151)

Cut-off	II	BII	UII
≤29/30	-0.660	0.022	-0.988
≤28/30	-0.606	0.048	-0.970
≤27/30	-0.476	0.124	-0.918
≤26/30	-0.328	0.212	-0.850
≤25/30	-0.166	0.306	-0.762
≤24/30	-0.010	0.392	-0.670
≤23/30	0.120	0.468	-0.572
≤22/30	0.250	0.536	-0.462
≤21/30	0.374	0.588	-0.348
≤20/30	0.476	0.618	-0.246
≤19/30	0.542	0.614	-0.184
≤18/30	0.608	0.600	-0.126
≤17/30	0.616	0.558	-0.146
≤16/30	0.682	0.582	-0.034
≤15/30	0.718	0.554	0.006
≤14/30	0.734	0.490	-0.010
≤13/30	0.732	0.416	-0.114
≤12/30	0.732	0.374	-0.164
≤11/30	0.724	0.304	-0.276

Abbreviations: II = identification index, BII = balanced identification index, UII = unbiased identification index

Figure 3: Plot of II (◆), BII (▲) and UII (×) values (y axis) vs MACE cut-off score (x-axis)