The International Journal of Biostatistics

Volume 5, Issue 1

2009

Article 11

Properties of the Projected Length of the Curve (PLC) and Area Swept out by the Curve (ASC) Indices for the Receiver Operating Characteristic (SROC) Curve

Xuan Zhang, *McMaster University*

Stephen D. Walter, *McMaster University*

Ramanakumar V. Agnihotram, *McGill University*

Recommended Citation:

Zhang, Xuan; Walter, Stephen D.; and Agnihotram, Ramanakumar V. (2009) "Properties of the Projected Length of the Curve (PLC) and Area Swept out by the Curve (ASC) Indices for the Receiver Operating Characteristic (SROC) Curve," *The International Journal of Biostatistics*: Vol. 5: Iss. 1, Article 11.

DOI: 10.2202/1557-4679.1096

Properties of the Projected Length of the Curve (PLC) and Area Swept out by the Curve (ASC) Indices for the Receiver Operating Characteristic (SROC) Curve

Xuan Zhang, Stephen D. Walter, and Ramanakumar V. Agnihotram

Abstract

Several measures have been proposed to summarize the Receiver Operating Characteristic (ROC) curve, including the Projected Length of the Curve (PLC) and the Area Swept out by the Curve (ASC). These indices were first proposed by Lee (Epidemiology 1996; 7:605-611) to avoid certain deficiencies of the traditional Area Under the Curve (AUC) summary measure. More recently meta-analysis methods for assessing diagnostic test accuracy have been developed and the Summary Receiver Operating Characteristic (SROC) curve has been recommended to represent the performance of a diagnostic test. Some properties of the SROC curve were discussed by Walter (Statist. Med. 2002; 21:1237-1256). Here we extend that work to focus on properties of PLC and ASC in the context of SROC curve. Mathematical expressions for these two indices and their variances are derived in terms of the overall diagnostic odds ratio and the magnitude of inter-study heterogeneity in the odds ratio. Expressions for PLC and ASC and their variances are easily computed in homogeneous studies, and their values provide good approximations to the corresponding values for heterogeneous studies in most practical situations. General variances of PLC and ASC are derived by using delta methods, and are found to be smaller if the odds ratio is large. The methods are illustrated using data from two studies, the first being a meta-analysis on the detection of metastases in cervical cancer patients, and the second being a single study of HPV infection and pre-invasive cervical lesions.

KEYWORDS: ROC curve, SROC curve, meta-analysis, the Projected Length of the Curve (PLC), the Area Swept out by the Curve (ASC)

Author Notes: We acknowledge the assistance of Dr. Eduardo Franco (McGill University) for the provision of the numerical data in our second example.

1. INTRODUCTION

The Receiver Operating Characteristic (ROC) curve has been recommended to represent the performance of a diagnostic test using data from a single study (Erdreich and Lee, 1981; Swets, 1988; Lee and Hsiao, 1996; Hanley, 1998; Walter, 2002; Zhou et al., 2002; Seong et al., 2004). The ROC curve shows the relationship between test sensitivity and specificity as the test threshold defining a positive result is varied. The Area Under the ROC curve (AUC) has been recognized as a useful index to provide a global picture of the test performance as the false-positive rate varies between 0 and 1 (Hanley and McNeil, 1982; Hilden, 1991; Zhou et al., 2002; Seong et al., 2004). Lee and Hsiao (1996) pointed out that AUC can sometimes erroneously suggest that a perfect diagnostic test has low accuracy. They therefore proposed two alternative summary indices, the Projected Length of the Curve (PLC) and the Area Swept out by the Curve (ASC). These two indices can avoid this shortcoming of the traditional AUC index.

More recently meta-analytic methods for evaluating diagnostic test accuracy have been developed. The Summary Receiver Operating Characteristic (SROC) curve was proposed as a means of summarizing the data on a test's sensitivity and specificity from multiple studies, while allowing for possibly different thresholds between studies (Irwig et al., 1993; Midgette et al., 1993; Moses et al., 1993; Walter, 2002; Zhou et al., 2002). New methods based on the SROC curve have since been developed, but little is known about the properties of the summary indices PLC and ASC expressions in this context.

In Section 2 and 3 some properties of the ROC and SROC curves are discussed. In Section 4 and 5 we derive explicit expressions for PLC and ASC and then evaluate them. In Section 6 and 7, the variances of PLC and ASC are derived for situations when studies are homogeneous or heterogeneous with respect to their diagnostic odds ratios. Two practical examples are discussed in Section 8. The first involves a meta-analysis of three diagnostic tests for metastases in cervical cancer patients and the second involves a single study of HPV infection in patients diagnosed with pre-invasive cervical lesions. Further points of discussion are considered in Section 9, including the performance of PLC and ASC in various situations when the magnitude of inter-study heterogeneity varies, and use of alternative models for meta-analysis of diagnostic test data.

2. ROC CURVE AND ITS INDICES: PLC AND ASC

The characteristic of primary interest for a diagnostic test is its ability to distinguish between diseased and non-diseased individuals at a particular test threshold (Zhou et al., 2002). For this purpose, measures such as the true positive

rate (TPR) and the false positive rate (FPR) are often considered. TPR / FPR (or sensitivity / [1-specificity]) is the likelihood ratio of probabilities that the test result is positive at a particular test threshold, given that the condition is present vs. absent. The relationship between TPR and FPR is represented by the ROC curve. The ROC curve for a single study is defined by one or more pairs of TPR and FPR values plotted in a unit square, and then connected empirically to graphically represent test performance for the whole range of possible thresholds. A schematic ROC curve is shown in Figure 1. Points on the empirical ROC curve are generated by considering all possible test thresholds. The smooth line is the corresponding fitted ROC curve.

Figure 1. A schematic empirical and fitted ROC curve. TPR = true positive rate; FPR = false positive rate.

Summary indices associated with the ROC curve can measure the overall accuracy of a test. The most familiar index is the Area Under the Curve (AUC) (Hanley and McNeil, 1982; Hilden, 1991; Walter, 2002; Zhou et al., 2002; Seong et al., 2004). Because the entire ROC curve is defined within a unit square, AUC ranges between 0 and 1. The closer AUC is to 1, the better the overall diagnostic performance of the test, and a test with AUC = 1 is perfectly accurate. For a test that discriminates disease cases and non-cases purely at random, the ROC curve would fall along the diagonal line from (0, 0) to (1, 1), and with AUC = 0.5. It has also been noted that AUC can also be interpreted as the probability that a randomly chosen test result for a case and a randomly chosen non-test result are correctly ranked in terms of suspicion of disease.

Although the AUC index plays an important role in summarizing the entire ROC curve, it is sometimes misleading when used as a measure of diagnostic accuracy. Lee and Hsiao (1996) and Zhou et al. (2002) presented examples of a perfect discriminating test, but with an AUC of only 0.5, the value corresponding to a completely random test. To avoid these drawbacks of AUC, Lee and Hsiao (1996) proposed two alternative summary indices for the ROC curve: PLC (the Projected Length of the Curve) and ASC (the Area Swept out by the Curve). Both these two indices represent the deviations of the ROC curve from the diagonal line and are, in some circumstances, better than AUC in representing the diagnostic accuracy of the test.

PLC is the sum (or integral) of all of the orthogonally projected lengths of the ROC curve onto the negative diagonal line (Lee and Hsiao, 1996). We use the ROC curve in Figure 2 (a) as a schematic example. The curve consists of three line segments (A, B and C), with the corresponding projected lengths a' , b' and c' , respectively. Hence the PLC for this example is $a' + b' + c'$.

For ASC, we imagine a ray emanating from the origin (0, 0) to a point on the ROC curve. As the point moves from the origin to the right-uppermost point (1, 1), the ray will sweep out some areas, and the total area swept out is denoted as ASC (Lee and Hsiao, 1996). We use the ROC curve in Figure 2 (b) as a schematic example. The ROC curve again consists of three line segments: A, B and C. The first ray emanating from the origin is segment A. As the end of the ray moves along segment B, an area $a1 + a2$ is swept out. Finally the end of the ray moves along line segment C, and an area $a2 + a3$ is swept out. Therefore, the ASC in this example is $a1 + 2 \times (a2) + a3$.

Figure 2. Two schematic ROC curves, each consists of three segments of straight lines to demonstrate (a) the calculation of PLC (b) the calculation of ASC.

For a test with no diagnostic value, the ROC curve lies entirely on the diagonal, and both PLC and ASC are zero. If the ROC curve deviates from the diagonal at any point, PLC and ASC become positive. The maximum value that PLC can attain is $\sqrt{2}$, and the corresponding value of ASC is $1/2$. In general, higher values of both PLC and ASC correspond to better test performance. In the extreme examples given by Lee and Hsiao (1996) and Zhou et al. (2002), PLC and ASC do indeed assume their maximum values and hence indicate strong discrimination between disease cases and non-cases; this is in contrast to the merely random test performance suggested by the use of AUC in these examples.

More recently meta-analysis statistical methods for synthesizing diagnostic test accuracy studies have been developed. In a meta-analysis the Summary Receiver Operating Characteristic (SROC) curve replaces the ROC curve as a way to summarize a test's TPR and FPR from multiple studies, while allowing the studies to have potentially different diagnostic thresholds for the test results (Irwig et al., 1993; Midgette et al., 1993; Moses et al., 1993; Walter, 2002; Zhou et al., 2002).

3. SROC CURVE AND ITS PROPERTIES

One way to analyze diagnostic test data from several studies is to plot TPR against FPR for each study, and then fit a SROC curve to these data points. For a

single study two measures are associated with the accuracy of a test: the Odds Ratio (OR) and the test threshold. OR summarizes the diagnostic performance of the test, and is given by

$$OR = \frac{TPR/(1-TPR)}{FPR/(1-FPR)} . \quad (1)$$

An implicit measure of the test threshold is

$$S = \ln\left(\frac{TPR}{1-TPR}\right) + \ln\left(\frac{FPR}{1-FPR}\right) . \quad (2)$$

If $TPR = 1 - FPR$, which implies equal false positive and false negative rates, then $S = 0$.

OR may increase or decrease with the threshold, leading to an asymmetric SROC curve. Letting $D = \ln(OR)$, Moses et al. (1993) suggested plotting D against S to check whether the test accuracy varies with the threshold. This approach transforms the observations from (FPR, TPR) space to (S, D) space. Under an assumption that the distributions of the test values for persons with and without the condition of interest follow a logistic distribution or, approximately, a normal distribution, the SROC curve becomes a straight line in (S, D) space. Accordingly, a linear relationship between D and S for all possible thresholds can be postulated through the following regression model:

$$D = a + bS. \quad (3)$$

If this straight line in (S, D) space is transformed back to SROC space, a smoothed fit to the SROC curve can be achieved (Moses et al., 1993). The results presented in this paper are based on the Moses model. Note that while this model was proposed in the context of data from multiple studies, it is also applicable to data from a single study under the assumption of logistic distributions for the test results of disease cases and non-cases (Walter and Sinuff, 2007). Accordingly, many of the properties of PLC and ASC that we will consider pertain to both the SROC and ROC curves.

The regression slope b represents the dependence of the test accuracy on threshold. When $b = 0$, the studies are homogeneous with respect to their diagnostic odds ratios, and therefore they can be summarized by an overall OR (Walter, 2002). In this case, the SROC curve is symmetric. When $b \neq 0$, the studies are heterogeneous with respect to OR (Walter, 2002), and OR varies between studies. Because test accuracy varies with the threshold, the resulting

SROC curve is asymmetric. The intercept a can be thought of as the value of $\ln(\text{OR})$ if S or b is 0. Under the regression model (3), the relationship between TPR and FPR can be expressed as

$$TPR = \frac{\exp(\frac{a}{1-b})(\frac{FPR}{1-FPR})^{\frac{1+b}{1-b}}}{1 + \exp(\frac{a}{1-b})(\frac{FPR}{1-FPR})^{\frac{1+b}{1-b}}} . \quad (4)$$

This expression gives TPR at any value of FPR, and hence defines the entire SROC curve. See Walter (2002) for a general review of the properties of this curve and further interpretation of its parameters.

Figure 3. SROC with various values of b ($a = 2$) under model (3)

Two special points on the SROC curve are of interest. First, there is a point Q^* on the SROC curve where $TPR = 1 - FPR$; the family of curves defined by a fixed value of a all pass through this point (see Figure 3). The point Q^* corresponds to the threshold at which the probability of a correct diagnosis is constant for all subjects, and has co-ordinates

$$TPR = \frac{\exp(a/2)}{1 + \exp(a/2)} \quad \text{and} \quad FPR = \frac{1}{1 + \exp(a/2)} .$$

When $b = 0$, the SROC curves are concave downwards. However, when $b \neq 0$, a portion (typically small) of the SROC curve lies below the diagonal,

and it has a small upwards concavity. In this region $TPR < FPR$ and the test would be performing worse than at random. This region is shown, for example, in the curve for $b = 0.5$ in Figure 3. We denote the second special point of interest, where the curve crosses the diagonal, as Q' , and it has co-ordinates

$$TPR = FPR = \frac{\exp(-a/2b)}{1 + \exp(-a/2b)}.$$

4 THE PROJECTED LENGTH OF THE CURVE (PLC)

In this section, we derive expressions for PLC, and consider its properties in various situations. As mentioned earlier, PLC is the sum of all of the projected lengths of the ROC curve onto the negative diagonal line. Under Moses' model we can express PLC in terms of the parameters a and b . To obtain an expression for PLC, we need to describe the form of the curves for various values of a and b . Figure 3 illustrates some possible behaviors of the curve. When $b \neq 0$, two separate regions of the curve lie below and above the diagonal. We denote the distance between the curve and the diagonal as d_1 (below the diagonal) and d_2 (above the diagonal); d_1^* and d_2^* are their corresponding maximum values, respectively. PLC is then defined as

$$PLC = 2(d_1^* + d_2^*). \quad (5)$$

Detailed expressions for d_1^* and d_2^* are derived in Appendix (A1).

Like AUC, PLC also has a probabilistic interpretation, as noted by Lee and Hsiao (1996). In single studies, PLC can be regarded as the probability of a correct diagnosis (as diseased or non-diseased) under the scenario whereby a subject has equal chances *a priori* of being truly diseased or non-diseased (Lee and Hsiao, 1996). This probability has been shown to equal $\frac{1}{2} + \frac{\sqrt{2}}{4} PLC$ (Lee and Hsiao, 1996), and we will denote it as P_{L1} . In the context of meta-analyses, this probability is

$$P_{L1} = \frac{1}{2} + \frac{\sqrt{2}}{4} PLC, \quad \text{when } a \geq 0, \quad (6)$$

$$P_{L1} = \frac{1}{2} - \frac{\sqrt{2}}{4} PLC, \quad \text{when } a < 0. \quad (7)$$

The maximum of PLC is $\sqrt{2}$. As $a \rightarrow \infty$, $PLC \rightarrow \sqrt{2}$ with a curve that is close to the ideal upper-left corner of the ROC space when $a > 0$. The corresponding P_{L1} is then $\frac{1}{2} + \frac{\sqrt{2}}{4} \times \sqrt{2} = 1$, so the test performs perfectly.

When $b = 0$, PLC is twice the distance between Q^* and point $O(\frac{1}{2}, \frac{1}{2})$, and thus we obtain

$$PLC_{\text{hom}} = \sqrt{8} \left| \frac{1}{2} - \frac{1}{1 + \exp(\frac{a}{2})} \right|, \quad (8)$$

where PLC_{hom} indicates the PLC for homogeneous studies. From equations (6)-(8), $P_{L1} = 1 - [1 + \exp(\frac{a}{2})]^{-1}$. Note that (8) implies PLC_{hom} is symmetric in a . In the special case $a = b = 0$, expression (8) degenerates to PLC_{hom} , which corresponds to the curve lying on the diagonal, indicating a test with only random diagnostic value. Since the expression for PLC_{hom} involves only the single parameter a , it is easy to compute PLC if the studies are homogeneous.

Table I. Values of (a) PLC and (b) PLC* for various values of a and b .

(a)				
$\begin{smallmatrix} a \\ \backslash \\ b \end{smallmatrix}$	0	1	2	3
0	0.000	0.346	0.654	0.898
0.05	0.063	0.350	0.655	0.899
0.10	0.127	0.360	0.659	0.902
0.15	0.190	0.379	0.667	0.906
0.20	0.254	0.409	0.678	0.911
0.25	0.318	0.446	0.693	0.919
0.30	0.382	0.490	0.714	0.929

(b)				
$\begin{smallmatrix} a \\ \backslash \\ b \end{smallmatrix}$	0	1	2	3
0	0.000	0.245	0.462	0.635
0.05	0.045	0.247	0.463	0.636
0.10	0.090	0.255	0.466	0.637
0.15	0.134	0.268	0.472	0.640
0.20	0.179	0.289	0.479	0.644
0.25	0.225	0.316	0.490	0.650
0.30	0.270	0.346	0.505	0.657

PLC* are the standardized values which are obtained by scaling PLC by $PLC_{\text{max}} = \sqrt{2}$.

For any given value of a , it turns out that PLC_{hom} is a useful lower bound for PLC in heterogeneous studies, as illustrated by our numerical evaluations. Table I illustrates the behaviour of PLC in heterogeneous and homogeneous cases,

and where the values of a (ranging from 0 to 3) and b (-0.3 to 0.3) cover most of the situations likely to occur in practice. Negative values of a are unlikely in practice because they correspond to the situation where the association of test result and disease status is in the opposite direction from that expected clinically. For a more convenient numerical comparison with AUC, we also standardized the original PLC values by dividing the original values by $PLC_{\max} = \sqrt{2}$; the standardized values then range between 0 and 1, as do the values of AUC. We denote the standardized PLC as PLC^* , and they are shown in Table I (b). For fixed a , PLC is symmetric with respect to b , so that the same summary value will be obtained for a given strength of dependence of diagnostic accuracy on the test threshold, regardless of its sign. PLC declines slowly as b increases from negative values to 0 and increases slowly as b increases from 0 to positive values. PLC is therefore minimized in the homogeneous case, when $b = 0$. For $|b| < 0.4$ and $|a| > 2$, the percentage difference in PLC compared to the homogeneous case is always less than 3 per cent. This indicates that PLC depends primarily on the value of a , and that PLC_{hom} provides a good approximation for heterogeneous studies in most practical situations.

5. THE AREA SWEEPED OUT BY THE CURVE (ASC)

We now derive expressions for ASC and consider its properties in various situations. ASC is the area swept out by the ROC curve, and it can also be expressed in terms of the parameters a and b . For illustration, we can imagine a ray emanating from the origin (0, 0) to each point on the curve. As the end of the ray moves from (0, 0) towards (1, 1), some areas are swept out within the ROC space. As shown in Figure 4, the area A is swept out first. As the ray moves upward along the curve, area C is swept out. When the ray moves through a point of tangency on the upper-left part of the curve, it folds back and areas C+B are then swept out. In this case, region C is swept out twice, and ASC is therefore defined as

$$ASC = A + B + 2C \quad (9)$$

(Lee and Hsiao, 1996). If we denote the area swept out as the ray moves from the origin up to the tangent line as M (the shaded area in Figure 4), area C can be expressed as

$$C = M - A, \quad \text{when } a > 0, b > 0, \quad (10)$$

$$C = M - B, \quad \text{when } a > 0, b < 0, \quad (11)$$

and area B can be expressed as

$$B = AUC + A - \frac{1}{2}, \quad \text{when } a > 0, \quad (12)$$

$$B = A + \frac{1}{2} - AUC, \quad \text{when } a < 0. \quad (13)$$

Here, by definition, AUC can be calculated as

$$AUC = \int_0^1 \frac{\exp(\frac{a}{1-b})(\frac{x}{1-x})^{\frac{1+b}{1-b}}}{1 + \exp(\frac{a}{1-b})(\frac{x}{1-x})^{\frac{1+b}{1-b}}} dx. \quad (14)$$

If we substitute expressions (10) – (13) in (9), after some simplification we can show that

$$ASC = AUC + 2M - \frac{1}{2}, \quad \text{when } b > 0 \text{ or} \quad (15)$$

$$ASC = 2M + \frac{1}{2} - AUC, \quad \text{when } b < 0. \quad (16)$$

Detailed expressions for ASC are derived in Appendix (A2).

Similarly to PLC, ASC is related to the probability of a correct diagnosis under the scenario whereby, among two randomly chosen diseased/non-diseased subjects, the one with higher test value (for example) is diagnosed as diseased when he/she is actually diseased, or equivalently the probability of being diagnosed as non-diseased when the one with higher test value is actually non-diseased (Lee and Hsiao, 1996). This is analogous to the probabilistic interpretation of the AUC; it can be interpreted as the probability that, for a randomly selected case/non-case pair of subjects, the diseased subject will have the more suspicious (case-like) test result (Lee and Hsiao, 1996; Walter, 2002; Zhou et al., 2002). This probability has been shown to equal $\frac{1}{2} + ASC$ (Lee and Hsiao, 1996), which we will denote as P_{L2} . Under the Moses model, this quantity becomes

$$P_{L2} = \frac{1}{2} + ASC, \quad \text{when } a \geq 0, \quad (17)$$

$$P_{L2} = \frac{1}{2} - ASC, \quad \text{when } a < 0. \quad (18)$$

If $a = 0$, $AUC = \frac{1}{2}$ and, from (15) or (16), $ASC = 2M$. If $b = 0$, no area is swept out twice and $M = 0$. ASC is then the region between the curve and the diagonal. Hence from (15) and (16)

$$ASC_{\text{hom}} = AUC_{\text{hom}} - \frac{1}{2}, \quad \text{when } a > 0 \text{ or} \quad (19)$$

$$ASC_{\text{hom}} = \frac{1}{2} - AUC_{\text{hom}}, \quad \text{when } a < 0, \quad (20)$$

where ASC_{hom} and AUC_{hom} indicate the ASC and AUC for homogeneous studies. From (14)

$$AUC_{\text{hom}} = \int_0^1 \frac{\exp(a)(\frac{x}{1-x})}{1 + \exp(a)(\frac{x}{1-x})} dx = \frac{\exp(a)}{[\exp(a) - 1]^2} [\exp(a) - 1 - a] \quad (21)$$

(Walter, 2002). Note that from equations (17) - (18) and equations (19) - (20), P_{L2} is equal to AUC_{hom} . If $a = b = 0$, the corresponding AUC is equal to $\frac{1}{2}$. Then from (19) or (20), $ASC_{\text{hom}} = 0$. Again, the zero value of ASC corresponds to the curve running along the chance diagonal. Once the curve deviates from the diagonal line, ASC will be strictly positive.

ASC does not necessarily attain its minimum as ASC_{hom} . In fact, the minimum occurs at various values of b , depending on the value of a . As we see in Figure 4, as a becomes large, ASC approaches its upper bound of 0.5, and the curve passes close to the ideal upper-left corner of the SROC space. The corresponding P_{L2} is then $\frac{1}{2} + \frac{1}{2} = 1$.

Figure 4. SROC curve to illustrate ASC under model (3)

Table II shows the behaviour of ASC for the heterogeneous and homogeneous cases. The original ASC values were standardized by dividing by $ASC_{\max} = 0.5$ so that the standardized values lie between 0 and 1, and we denote the standardized ASC as ASC^* . For fixed a , ASC declines as b increases from negative values to 0 and increases as b increases from 0 to positive values. For

small values of a , ASC depends more strongly on b and takes small values when b is close to 0. ASC is minimized when $a = 0$. For fixed a , ASC depends only weakly on b . For $|b| < 0.4$ and $|a| > 1$, the percentage difference in ASC compared to the homogeneous case is less than 2 per cent. Thus, as was seen for PLC, ASC_{hom} provides a good approximation in heterogeneous studies.

Table II. Values of (a) ASC and (b) ASC* for various values of a and b .

(a)				
$\begin{smallmatrix} a \\ b \end{smallmatrix}$	0	1	2	3
-0.30	0.148	0.195	0.295	0.381
-0.25	0.123	0.182	0.293	0.382
-0.20	0.099	0.171	0.292	0.384
-0.15	0.074	0.164	0.293	0.385
-0.10	0.049	0.161	0.294	0.386
-0.05	0.025	0.161	0.294	0.387
0	0.000	0.161	0.294	0.387
0.05	0.025	0.161	0.294	0.387
0.10	0.049	0.163	0.294	0.386
0.15	0.074	0.169	0.295	0.386
0.20	0.099	0.178	0.296	0.385
0.25	0.123	0.190	0.299	0.385
0.30	0.148	0.205	0.304	0.386

(b)				
$\begin{smallmatrix} a \\ b \end{smallmatrix}$	0	1	2	3
-0.30	0.295	0.391	0.591	0.762
-0.25	0.246	0.363	0.586	0.764
-0.20	0.197	0.342	0.585	0.767
-0.15	0.148	0.328	0.586	0.770
-0.10	0.099	0.323	0.587	0.772
-0.05	0.049	0.322	0.589	0.774
0	0.000	0.323	0.589	0.774
0.05	0.049	0.323	0.589	0.774
0.10	0.099	0.327	0.588	0.773
0.15	0.148	0.338	0.589	0.771
0.20	0.197	0.356	0.592	0.770
0.25	0.246	0.380	0.598	0.770
0.30	0.295	0.409	0.607	0.771

ASC* are the standardized values which are obtained by scaling ASC by $ASC_{\text{max}} = 0.5$.

6. THE SAMPLE VARIANCE OF THE ESTIMATED PLC

We may derive approximate variances for the estimates of PLC and ASC using the delta method. For PLC (see Appendix A1), we see that PLC is a function of the regression parameters a and b . Hence, an approximate variance for $\hat{PLC}$ is

$$\text{var}(\hat{PLC}) = \left(\frac{\partial PLC}{\partial a} \right)^2 \text{var}(\hat{a}) + \left(\frac{\partial PLC}{\partial b} \right)^2 \text{var}(\hat{b}) + 2 \left(\frac{\partial PLC}{\partial a} \right) \left(\frac{\partial PLC}{\partial b} \right) \text{cov}(\hat{a}, \hat{b}) \quad (22)$$

where, from (5)

$$\frac{\partial PLC}{\partial a} = 2 \frac{\partial d_1^*}{\partial a} + 2 \frac{\partial d_2^*}{\partial a}, \quad (23)$$

and

$$\frac{\partial PLC}{\partial b} = 2 \frac{\partial d_1^*}{\partial b} + 2 \frac{\partial d_2^*}{\partial b}. \quad (24)$$

The expansions of the partial derivatives (23) and (24) can be found in Appendix (A3). In the case of homogeneous studies where $b = 0$, equation (22) becomes

$$\text{var}(\hat{PLC}) \approx \left(\frac{\partial PLC_{\text{hom}}}{\partial a} \right)^2 \text{var}(\hat{a}),$$

which leads to (see Appendix (A5))

$$\text{var}(\hat{PLC}_{\text{hom}}) = \frac{2 \exp(a)}{[1 + \exp(\frac{a}{2})]^4} \text{var}(\hat{a}). \quad (25)$$

As suggested by (8), (25) implies that $\text{var}(\hat{PLC}_{\text{hom}})$ is symmetric in a . When $a = 0$, $\text{var}(\hat{PLC}_{\text{hom}}) = \frac{1}{8} \text{var}(\hat{a})$.

To illustrate the empirical performance of $\text{var}(\hat{PLC})$ in homogeneous and heterogeneous studies, we use expression (22) with the observed values of $\text{var}(\hat{a})$, $\text{var}(\hat{b})$ and $\text{cov}(\hat{a}, \hat{b})$ based on data from a typical example, involving the lymphangiography (LAG) test for the diagnosis of lymph node metastasis in cervical cancer patients (Scheidler et al., 1997). These values are: 0.1430, 0.0631 and 0.0529 respectively. In Table III, it is evident that the heterogeneous variances are larger than the homogeneous estimates. $\text{var}(\hat{PLC})$ is minimized when $b = 0$, which is therefore the best situation for estimating PLC. For fixed b , as $|a| \rightarrow \infty$, the SROC curve moves closer to the upper-left corner, and $\text{var}(\hat{PLC})$ decreases; this would indicate a perfect test. For fixed b , as $|a| \rightarrow 0$,

the SROC curve moves closer to the diagonal line, $\text{var}(\hat{PLC})$ increases and is maximized when $a = 0$; this indicates the worst situation for estimating PLC, when the diagnostic test has close to random performance. For large values of a , the variance for the homogeneous case provides a good approximation to heterogeneous studies, with a difference of less than 1 per cent. As expected, $\text{var}(\hat{PLC})$ is symmetric in both a and b .

Table III. Results from the LAG test for the diagnosis of lymph node metastasis in cervical cancer patients: (a) values of $\text{var}(\hat{PLC})$ and (b) values of $\text{var}(\hat{PLC}^*)$ for various values of a and b

(a)				
$\begin{smallmatrix} a \\ b \end{smallmatrix}$	0	1	2	3
0	0.018	0.016	0.011	0.006
0.05	0.101	0.021	0.013	0.007
0.10	0.101	0.029	0.015	0.008
0.15	0.102	0.042	0.018	0.009
0.20	0.102	0.056	0.021	0.010
0.25	0.103	0.069	0.026	0.011
0.30	0.104	0.078	0.032	0.013

(b)				
$\begin{smallmatrix} a \\ b \end{smallmatrix}$	0	1	2	3
0	0.009	0.008	0.006	0.003
0.05	0.051	0.011	0.007	0.004
0.10	0.051	0.015	0.008	0.004
0.15	0.051	0.021	0.009	0.005
0.20	0.051	0.028	0.011	0.005
0.25	0.052	0.035	0.013	0.006
0.30	0.052	0.039	0.016	0.007

7. THE SAMPLE VARIANCE OF THE ESTIMATED ASC

The delta method yields an approximate variance for ASC as

$$\text{var}(\hat{ASC}) = \left(\frac{\partial ASC}{\partial a} \right)^2 \text{var}(\hat{a}) + \left(\frac{\partial ASC}{\partial b} \right)^2 \text{var}(\hat{b}) + 2 \left(\frac{\partial ASC}{\partial a} \right) \left(\frac{\partial ASC}{\partial b} \right) \text{cov}(\hat{a}, \hat{b}) \quad (26)$$

where,

$$\frac{\partial ASC}{\partial a} = \frac{\partial AUC}{\partial a} + 2 \frac{\partial M}{\partial a}, \quad \text{when } b > 0, \quad \text{or} \quad (27)$$

$$\frac{\partial ASC}{\partial a} = 2 \frac{\partial M}{\partial a} - \frac{\partial AUC}{\partial a}, \quad \text{when } b < 0. \quad (28)$$

Similarly,

$$\frac{\partial ASC}{\partial b} = \frac{\partial AUC}{\partial b} + 2 \frac{\partial M}{\partial b}, \quad \text{when } b > 0, \quad \text{or} \quad (29)$$

$$\frac{\partial ASC}{\partial b} = 2 \frac{\partial M}{\partial b} - \frac{\partial AUC}{\partial b}, \quad \text{when } b < 0. \quad (30)$$

Expansions of the partial derivatives (27) – (30) can be found in Appendix (A4). When $a = 0$, $ASC = 2M$. Thus, $\frac{\partial ASC}{\partial b} = 2 \frac{\partial M}{\partial b}$, and the approximate variance for ASC then degenerates to

$$\text{var}(A\hat{S}C) = 4 \left(\frac{\partial M}{\partial b} \right)^2 \text{var}(\hat{b}), \quad (31)$$

which is symmetric in b .

In the case of homogeneous studies where $b = 0$, from (19) and (20), the delta method gives an approximate variance for ASC_{hom} as

$$\text{var}(A\hat{S}C) = \left(\frac{\partial AUC_{\text{hom}}}{\partial a} \right)^2 \text{var}(\hat{a}) = \text{var}(A\hat{U}C_{\text{hom}}).$$

As derived in Appendix (A6), we may obtain

$$\text{var}(A\hat{S}C_{\text{hom}}) = \frac{\exp(2a)}{[\exp(a) - 1]^6} \{a[\exp(a) + 1] - 2[\exp(a) - 1]\}^2 \text{var}(\hat{a}). \quad (32)$$

In the special case $a = b = 0$, expression (32) is degenerate, but by using L'Hopital's rule one can show that here $\text{var}(A\hat{S}C_{\text{hom}}) \approx \frac{1}{6} \text{var}(\hat{a})$.

To illustrate the behavior of $\text{var}(A\hat{S}C)$ we again use data from the lymphangiography (LAG) test (Scheidler et al., 1997). Table IV gives the values

of $\text{var}(\hat{ASC})$ in the homogeneous and heterogeneous cases. As seen from equations (31) and (32), $\text{var}(\hat{ASC})$ is symmetric in b when $a = 0$, and symmetric in a when $b = 0$. When $b \neq 0$, for fixed b , as $|a| \rightarrow 0$, $\text{var}(\hat{ASC})$ increases. $\text{var}(\hat{ASC})$ is maximized when $a = 0$, the worst situation for estimating ASC. The variance for the homogeneous case again appears to provide a good approximation for heterogeneous studies, with the percentage difference being less than 1 per cent.

Table IV. Results from the LAG test for the diagnosis of lymph node metastasis in cervical cancer patients: (a) values of $\text{var}(\hat{ASC})$ and (b) values of $\text{var}(\hat{ASC}^*)$ for various values of a and b

(a)				
$\begin{smallmatrix} a \\ b \end{smallmatrix}$	0	1	2	3
-0.30	0.0150	0.0094	0.0024	0.0006
-0.25	0.0151	0.0081	0.0020	0.0006
-0.20	0.0152	0.0064	0.0017	0.0006
-0.15	0.0153	0.0047	0.0016	0.0006
-0.10	0.0154	0.0033	0.0017	0.0006
-0.05	0.0154	0.0032	0.0017	0.0007
0	0.0040	0.0033	0.0018	0.0008
0.05	0.0154	0.0034	0.0018	0.0007
0.10	0.0154	0.0044	0.0018	0.0007
0.15	0.0153	0.0060	0.0020	0.0007
0.20	0.0152	0.0078	0.0023	0.0007
0.25	0.0151	0.0092	0.0027	0.0007
0.30	0.0150	0.0103	0.0031	0.0008

(b)				
$\begin{smallmatrix} a \\ b \end{smallmatrix}$	0	1	2	3
-0.30	0.0599	0.0377	0.0095	0.0024
-0.25	0.0604	0.0324	0.0079	0.0023
-0.20	0.0609	0.0257	0.0069	0.0023
-0.15	0.0612	0.0187	0.0065	0.0024
-0.10	0.0615	0.0139	0.0066	0.0026
-0.05	0.0616	0.0127	0.0070	0.0028
0	0.0159	0.0130	0.0073	0.0030
0.05	0.0616	0.0138	0.0071	0.0028
0.10	0.0615	0.0175	0.0073	0.0027
0.15	0.0612	0.0240	0.0079	0.0026
0.20	0.0609	0.0310	0.0090	0.0027
0.25	0.0604	0.0369	0.0106	0.0029
0.30	0.0599	0.0413	0.0126	0.0032

8. NUMERICAL EXAMPLES

To illustrate the methods, we consider the three summary measures, AUC, PLC and ASC in two examples, in the context of the SROC curve in the first and in the context of the ROC curve in the second. The first example concerns a meta-analysis of three imaging tests for the diagnosis of lymph node metastasis in patients with cervical cancer: lymphangiography (LAG), computed tomography (CT), and magnetic resonance (MR) (Scheidler et al., 1997). The data come from 17 studies that met the inclusion criteria for LAG, 19 for CT, and 10 for MR. The fitted SROC curves are plotted in Figure 5.

Table V gives the results for the three tests. The b values (or the degree of heterogeneity) are not equal to zero. Only the LAG test has a negative regression slope b , and consequently its SROC curve has a small region lying below the chance diagonal near the upper-right corner. As anticipated from the theoretical results, the homogeneous approximations are quite accurate in all cases. The probabilities P_{L1} and P_{L2} in the homogeneous and heterogeneous cases are bracketed below their corresponding PLC and ASC values. In this example, they all agree very closely with the numerical values of AUC.

Figure 5. SROC analysis of three tests for cervical cancer metastases: (a) LAG test (b) CT test (c) MR test

The original analysis by Scheidler et al. (1997) showed no significant differences in the overall performance of the three tests, but there was a trend toward better performance for MR imaging. The findings in Table V are compatible with Scheidler's results. Both AUC and the probabilities of a correct

diagnosis rank the tests in the same order of performance: MR imaging, CT and LAG. Also, the standard errors are relatively large for all three summary measures with the LAG data.

To facilitate a fair comparison of PLC* and ASC* with the AUC results, we should also define a rescaled AUC* value, taking values 0 for a random test and 1 for a perfect test; accordingly we use $AUC^* = 2(AUC - 0.5)$.

For the data in the first example, we obtain $AUC^* = 0.590, 0.734$ and 0.828 for the LAG, CT and MR tests respectively. These values exceed the corresponding values of PLC*, but are slightly inferior to the corresponding values of ASC*.

Table V. Results of SROC analysis of three tests for cervical cancer metastases.

(a) unscaled measures

	a	b	AUC	AUC_{hom}	PLC	PLC_{hom}	ASC	ASC_{hom}
LAG test	2.09	-0.35	0.795	0.805	0.758 (0.768)	0.678 (0.740)	0.308 (0.808)	0.305 (0.805)
(SE)*	0.38	0.25	0.049	0.041	0.188	0.103	0.051	0.041
CT test	2.79	0.22	0.867	0.871	0.870 (0.808)	0.852 (0.801)	0.370 (0.870)	0.371 (0.871)
(SE)*	0.36	0.12	0.026	0.029	0.094	0.081	0.028	0.029
MR test	3.51	0.25	0.914	0.919	1.013 (0.858)	0.997 (0.853)	0.416 (0.916)	0.419 (0.919)
(SE)*	0.61	0.19	0.028	0.033	0.123	0.108	0.031	0.033

SE refers to standard errors for the sample estimates of AUC, PLC and ASC in the homogeneous and heterogeneous cases. The bracketed numbers are p_{L1} and p_{L2} related to PLC and ASC in the homogeneous and heterogeneous cases.

(b) standardized measures

	a	b	AUC^*	AUC_{hom}^*	PLC^*	PLC_{hom}^*	ASC^*	ASC_{hom}^*
LAG test	2.09	-0.35	0.590	0.610	0.536	0.480	0.616	0.609
(SE)*	0.38	0.25	0.098	0.082	0.133	0.073	0.102	0.082
CT test	2.79	0.22	0.734	0.742	0.615	0.603	0.739	0.742
(SE)*	0.36	0.12	0.052	0.058	0.066	0.057	0.056	0.058
MR test	3.51	0.25	0.828	0.838	0.716	0.705	0.833	0.839
(SE)*	0.61	0.19	0.056	0.066	0.087	0.076	0.062	0.066

* SE refers to standard errors for the sample estimates of AUC^* , PLC^* and ASC^* in the homogeneous and heterogeneous cases.

The second example comes from a Brazilian longitudinal cohort study of HPV infection and pre-invasive cervical lesions on 2462 women (Franco et al, 1999). The Normalized Absorbance Ratio (NAR) is the numerical score for the antibody test that reveals exposure to HPV. It is continuous, with higher values indicating a higher likelihood of exposure to HPV. In this study NAR was

determined at the time of patient enrolment. Patients were classified into three diagnostic groups: (1) low squamous intraepithelial lesion at enrolment (lsil), (2) high squamous intraepithelial lesion at enrolment (hsil) and (3) high squamous intraepithelial lesion after one year follow-up (hsil_1). We again adopt Moses' model (3); while this model was proposed for meta-analysis and the fitting of SROC curves, note that it is also applicable to individual-level data from single studies.

Table VI. Frequency and percentage distributions of cervical lesions by various threshold values of diagnostic test (NAR) results.

Case definition: lsil	>3	>2.5	>2	>1.5	>1	>0.5	>0
Cases							
Number	0	0	4	5	6	9	5
Percentage rated equal or higher	0	0	12.9	29.0	48.4	83.9	100.0
Non-cases							
Number	32	54	79	227	566	1067	309
Percentage rated equal or higher	1.4	3.7	7.1	16.8	41.0	86.4	100.0

Case definition: hsil	>3	>2.5	>2	>1.5	>1	>0.5	>0
Cases							
Number	2	1	1	4	1	8	3
Percentage rated equal or higher	10.0	15.0	20.0	40.0	45.0	85.0	100.0
Non-cases							
Number	32	54	79	227	566	1067	309
Percentage rated equal or higher	1.4	3.7	7.1	16.8	41.0	86.4	100.0

Case definition: hsil_1	>3	>2.5	>2	>1.5	>1	>0.5	>0
Cases							
Number	4	1	1	5	2	14	4
Percentage rated equal or higher	12.9	16.1	19.4	35.5	41.9	87.1	100.0
Non-cases							
Number	33	54	83	235	580	1085	318
Percentage rated equal or higher	1.4	3.6	7.1	17.0	41.2	86.7	100.0

Table VI presents the data on the rating of the NAR values using three alternative threshold definitions for cervical lesions: lsil, hsil and hsil_1. The test threshold values are defined as: >3, >2.5, >2, >1.5, >1, >0.5 and >0. The fitted ROC curves are plotted in Figure 6, with each point on the curves representing a

(FPR, TPR) pair corresponding to a different test threshold. The coordinates of these points are presented in Table VI.

Figure 6. ROC curves for Brazilian longitudinal cohort study of HPV diagnostic test (NAR) and cervical neoplasia defined as (a) lsil, (b) hsil, and (c) hsil-1

The values of the AUC, PLC and ASC indices and the probabilities P_{L1} and P_{L2} related to PLC and ASC in these three curves are shown in Table VII. For this example, both PLC* and ASC* are substantially higher than AUC* for all three diagnostic thresholds. The standardized values PLC* and ASC* are approximately 50% greater than the corresponding values of AUC* in all cases. P_{L1} and P_{L2} are approximately 0.6 for all three thresholds. Collectively these results suggest that PLC and ASC have somewhat higher diagnostic ability than AUC for these data. Recall however that the diagnostic scenarios are different for each summary measure, so such comparisons should be made cautiously, but we can note that there is some stability in the findings given that the results for all three measures did not vary across the three choices for the HPV diagnostic threshold.

Table VII. Values of AUC, PLC and ASC for the HPV diagnostic test – cervical neoplasia example

	AUC	PLC	ASC	P_{L1}	P_{L2}	AUC*	PLC*	ASC*
lsil (SE)*	0.540 0.004	0.173 0.029	0.068 0.020	0.561	0.568	0.080 0.008	0.122 0.021	0.136 0.040
hsil (SE)*	0.581 0.010	0.326 0.059	0.128 0.049	0.615	0.628	0.161 0.020	0.231 0.042	0.256 0.098
hsil_1 (SE)*	0.577 0.006	0.316 0.041	0.124 0.034	0.612	0.624	0.154 0.012	0.224 0.029	0.248 0.068

* SE refers to standard errors for the sample estimates of AUC, PLC, ASC, AUC*, PLC* and ASC* in the homogeneous and heterogeneous cases

9. DISCUSSION

In this paper we have established basic properties of PLC and ASC in the context of ROC and SROC curves. The mathematical expressions of these two indices and their variances were derived. We showed that the value of PLC in homogeneous studies is a reasonable lower bound approximation for PLC in heterogeneous studies. Similarly, ASC in the homogeneous case provides a good approximation to heterogeneous studies with a large odds ratio. PLC and ASC and their variances are easily computed in the homogeneous cases. The variances of PLC and ASC were found to be smaller if the odds ratio is large.

PLC and ASC, firstly proposed by Lee and Hsiao (1996) in the context of ROC curves for single studies, are said to avoid certain deficiencies of the traditional AUC summary index. Lee and Hsiao (1996) presented two hypothetical examples to show the drawbacks of the AUC index and to illustrate how PLC and ASC can avoid certain shortcoming and better evaluate a diagnostic test in these examples (Lee and Hsiao, 1996). In these examples, use of AUC would suggest only random accuracy of the diagnostic test, while use of PLC and ASC would indicate perfect discrimination between disease cases and non-cases. We used two examples based on real data. In the first numerical example in Section 8, we showed how the two indices are calculated and some of their properties, but the conclusions reached by using PLC, ASC and AUC were similar. In the second example, PLC and AUC had a numerical advantage over AUC when all three measures were standardized and expressed on a common scale from 0 (random test) to 1 (perfect test).

We noted that, for fixed a , negative values of b yield the same value of PLC as the equivalent positive values. Thus, regardless of the sign, the same summary value will be obtained for a given strength of dependence of diagnostic accuracy on the test threshold. However, although the values of ASC are not completely

symmetric in b , we find in our tabulations that for positive and negative values of b of the same magnitude, ASC differs by less than 1 per cent.

Use of a partial ROC space has been suggested if one wishes to represent test performance on those portions of the curve of clinical interest, or to exclude the parts where study data are sparse (Thompson and Zucchini, 1989; Walter, 2005). Further work is needed to explore the properties of the PLC and ASC in the context of the partial ROC space. Intuitively we may expect that the partial PLC lacks the symmetry enjoyed by the full PLC. For example, if we focus on the values of $FPR < 0.2$ when $a = 2$, the corresponding partial PLC will be greater when $b < 0$ than when $b > 0$.

Another area for future study is the finite sample behaviour of PLC and ASC and their variances. Empirical distributions of the parameters a and b could be simulated based on the regression equation (3), and then the behaviour of PLC and ASC and their variances could be compared to the approximate (asymptotic) formulae derived in this paper. While we have elected not to pursue this topic in detail in the present paper, we did briefly consider it for the more established summary measure AUC. We think that our approximate variances for PLC and ASC will in fact be adequate for the sample sizes typically encountered in analysis problems of the type we have described. Our intuition is that the asymptotic approximations will suffice for sample sizes similar to those when the binomial distribution is reasonably well approximated by the normal distribution. (Also recall here that PLC and ASC do have probabilistic interpretations.) Our belief is supported by Zhou et al. (2002) who comment that there must be “at least 30 subjects with and without the condition (disease) for the asymptotic intervals...to provide adequate coverage”. Further insight is provided by Pepe (2003), who remarks that the asymptotic variance for the empirical AUC provides a reasonable approximation when terms of the order n^{-2} are ignored. This is similar to ignoring second order terms in the delta method as we have applied it.

All the results presented here are based on Moses’ model being valid. The Moses method models the logit difference for TPR and FPR as a function of the logit sum for TPR and FPR. When $TPR = 0$ or $FPR = 0$ for a given study, the transformation is undefined. To avoid this problem, an adjustment is made by adding 0.5 to all cells of the 2×2 data table for that study (Walter, 1985; Moses et al., 1993). This adjustment is often applied to all studies in a meta-analysis and is asymptotically negligible, but it has been shown to result in a downward bias on estimated SROC curves (Macaskill, 2004). More research is needed to explore the feasibility of fitting the Moses model with a small number of studies. As noted, while no assumptions are made about the distribution of the independent variable S , Moses’ approach relies on asymptotic normality in the dependent variable D (Macaskill, 2004). Thus, the adequacy of the normal approximation for the estimates of D in small samples should be studied.

Moses' regression model only allows for a fixed effect to take the variability between studies into account. This implicitly assumes that the variability between studies is due only to sampling error. Equal weights are applied to studies fitting an ordinary least squares regression model, so that relatively higher weight is given to smaller studies compared to a random effects analysis.

While only fixed effects are permitted in Moses model, an alternative approach proposed by Rutter and Gatsonis (1995 and 2001) allows for fixed or random effects of the test threshold and accuracy. Because each study is assumed to have its own accuracy, this approach takes both the within- and between-variances into account. These within-study random effects are generally assumed to have a normal distribution. Because it is necessary to check the association between test accuracy and covariate information for heterogeneous studies, incorporation of random effects in the analysis of variability is often adopted in practice. It is preferable to analyze the data at the level of individual subjects, in order to better reflect the effects of individual level covariates and avoid potential ecologic bias associated with analyses of data aggregated to the study level. However, because of the limitations of current meta-analyses, and in particular the fact that data are seldom reported at this level, it is often not possible to examine the random effects model in this way.

Another alternative approach for fitting SROC curves is the Hierarchical SROC (HSROC) model proposed by Macaskill (2004). This model carries out two-level analysis, corresponding to within- and between- studies. Three parameters are involved in the model. Two of them are specified as random effects: one represents an implicit threshold; the other the diagnostic accuracy. These two parameters are assumed to be independent and normally distributed. The third parameter, which is assumed to be a fixed effect, provides asymmetry in the SROC by allowing test accuracy to vary with the threshold. In using this model, one would ideally not only obtain the summary estimates of sensitivity and specificity from each study, but also have access to the individual level data. In contrast, the Moses model allows only for a fixed effect, and so information from the individual level data is lost. However the Moses model has an advantage that it has only two parameters and it makes no assumption about the distribution of S (Macaskill, 2004). A recent commentary (Begg, 2008) has discussed other relative strengths and weaknesses of these various approaches to the meta-analysis of diagnostic test data.

APPENDIX

A1. Derivation of PLC

Since the curve can be expressed as $y = f(x, a, b) = \frac{\exp(\frac{a}{1-b})(\frac{x}{1-x})^{\frac{1+b}{1-b}}}{1 + \exp(\frac{a}{1-b})(\frac{x}{1-x})^{\frac{1+b}{1-b}}}$ and

the co-ordinates of point Q' are $TPR = FPR = \frac{\exp(-a/2b)}{1 + \exp(-a/2b)}$, by using the formula

for the distance from a point to a straight line, the maximum distance d_1^* and d_2^* can be expressed as

$$d_1^* = \frac{\hat{x}_1 - 1 + \frac{1}{1 + \exp(\frac{a}{1-b})(\frac{\hat{x}_1}{1-\hat{x}_1})^{\frac{1+b}{1-b}}}}{\sqrt{2}}, \quad (A1)$$

when $b > 0$, $0 < x < \frac{\exp(-a/2b)}{1 + \exp(-a/2b)}$; or when $b < 0$, $\frac{\exp(-a/2b)}{1 + \exp(-a/2b)} < x < 1$, and

$$d_2^* = \frac{-\hat{x}_2 + 1 - \frac{1}{1 + \exp(\frac{a}{1-b})(\frac{\hat{x}_2}{1-\hat{x}_2})^{\frac{1+b}{1-b}}}}{\sqrt{2}}, \quad (A2)$$

when $b > 0$, $\frac{\exp(-a/2b)}{1 + \exp(-a/2b)} < x < 1$; or when $b < 0$, $0 < x < \frac{\exp(-a/2b)}{1 + \exp(-a/2b)}$.

Here the distance d_1 has its maximum d_1^* at $\hat{x}_1$ and d_2 has the maximum d_2^* at $\hat{x}_2$. We could apply an iterative procedure to obtain the points $\hat{x}_1$ and $\hat{x}_2$, by the Newton-Raphson method, for example.

A.2 Derivation of ASC

The ray emanating from the origin (0,0) is the line $y = kx$. Letting $kx = f(x, a, b)$, we have

$$k = \frac{1}{x} f(x, a, b) = \frac{1}{x} \left[1 - \frac{1}{1 + \exp\left(\frac{a}{1-b}\right)\left(\frac{x}{1-x}\right)^{\frac{1+b}{1-b}}}\right]. \quad (\text{A3})$$

Differentiating both sides of equation (A3) with respect to x and equating to 0, after some simplification, we obtain

$$\frac{(1+b)}{(1-b)} \left(\frac{1}{1-x}\right) = 1 + \exp\left(\frac{a}{1-b}\right) \left(\frac{x}{1-x}\right)^{\frac{1+b}{1-b}}. \quad (\text{A4})$$

Denote $x = h(a, b)$ as the solution of (A4). The extreme value of k is

$$k^* = \frac{1}{h(a, b)} f(h(a, b), a, b) \quad (\text{A5})$$

and the point of tangency has co-ordinates $h(a, b)$ and $f(h(a, b), a, b)$. Lacking a closed-form solution, we have to apply an iterative root-finding technique in the determination of $h(a, b)$; again, the Newton-Raphson iteration method can be used.

Now ASC can be expressed as

$$\begin{aligned} ASC &= AUC + 2M - \frac{1}{2} \\ &= \int_0^1 f(x, a, b) dx + 2 \int_0^{h(a, b)} [k^* x - f(x, a, b)] dx - \frac{1}{2} \quad \text{when } b > 0 \end{aligned} \quad (\text{A6})$$

or

$$\begin{aligned} ASC &= 2M + \frac{1}{2} - AUC \\ &= 2 \int_0^{h(a, b)} [f(x, a, b) - k^* x] dx + \frac{1}{2} - \int_0^1 f(x, a, b) dx \quad \text{when } b < 0. \end{aligned} \quad (\text{A7})$$

A.3 Derivation of the variance of $\hat{PLC}$

To solve $\frac{\partial PLC}{\partial a}$, we need to find expressions for $\frac{\partial d_1^*}{\partial a}$ and $\frac{\partial d_2^*}{\partial a}$. From equation (A1), if we let $v_1(a, b) = \hat{x}_1$, we may derive

$$\frac{\partial d_1^*}{\partial a} = \frac{1}{\sqrt{2}} \left[\frac{\partial v_1(a, b)}{\partial a} - \frac{\frac{1}{1-b} \exp(\frac{a}{1-b}) (\frac{v_1(a, b)}{1-v_1(a, b)})^{\frac{1+b}{1-b}} \left[1 + \frac{(1+b)}{v_1(a, b)(1-v_1(a, b))} \frac{\partial v_1(a, b)}{\partial a} \right]}{[1 + \exp(\frac{a}{1-b}) (\frac{v_1(a, b)}{1-v_1(a, b)})^{\frac{1+b}{1-b}}]^2} \right] \quad (A8)$$

where

$$\frac{\partial v_1(a, b)}{\partial a} = \frac{\frac{1}{1-b} - 2[1 + \exp(\frac{a}{1-b}) (\frac{v_1(a, b)}{1-v_1(a, b)})^{\frac{1+b}{1-b}}] \frac{v_1(a, b)(1-v_1(a, b))}{1+b}}{2[1 + \exp(\frac{a}{1-b}) (\frac{v_1(a, b)}{1-v_1(a, b)})^{\frac{1+b}{1-b}}] - \frac{2b}{(1-b)v_1(a, b)(1-v_1(a, b))} - \frac{2}{1-v_1(a, b)}}. \quad (A9)$$

From equation (A2), if we let $v_2(a, b) = \hat{x}_2$, substitute $v_2(a, b)$ for $v_1(a, b)$ in equation (A8) and (A9), and also change the sign of equation (A8), we may obtain $\frac{\partial d_2^*}{\partial a}$.

Similarly, to solve $\frac{\partial PLC}{\partial b}$, we need to find expressions for $\frac{\partial d_1^*}{\partial b}$ and $\frac{\partial d_2^*}{\partial b}$.

Using the same methods, we may show after some simplification that $\frac{\partial d_1^*}{\partial b}$ is equal to

$$\frac{1}{\sqrt{2}} \left[\frac{\partial v_1(a, b)}{\partial b} - \frac{\frac{\exp(\frac{a}{1-b}) (\frac{v_1(a, b)}{1-v_1(a, b)})^{\frac{1+b}{1-b}}}{(1-b)^2} \left[a + 2 \ln(\frac{v_1(a, b)}{1-v_1(a, b)}) + \frac{(1-b^2)}{v_1(a, b)(1-v_1(a, b))} \frac{\partial v_1(a, b)}{\partial b} \right]}{[1 + \exp(\frac{a}{1-b}) (\frac{v_1(a, b)}{1-v_1(a, b)})^{\frac{1+b}{1-b}}]^2} \right] \quad (A10)$$

where $\frac{\partial v_1(a, b)}{\partial b}$ is equal to

$$\frac{\left\{ -\frac{1}{2(1-b)^2} + [1 + \exp(\frac{a}{1-b}) (\frac{v_1(a, b)}{1-v_1(a, b)})^{\frac{1+b}{1-b}}] \frac{v_1(a, b)(1+v_1(a, b))}{(1+b)^2} \right\} [a + 2 \ln(\frac{v_1(a, b)}{1-v_1(a, b)})] + \frac{1}{(1+b)^2}}{-[1 + \exp(\frac{a}{1-b}) (\frac{v_1(a, b)}{1-v_1(a, b)})^{\frac{1+b}{1-b}}] + \frac{b}{(1-b)v_1(a, b)(1+v_1(a, b))} + \frac{1}{1-v_1(a, b)}}. \quad (A11)$$

Similarly, changing $v_1(a, b)$ to $v_2(a, b)$ in equation (A10) and (A11), and changing the sign of equation (A10), we may derive $\frac{\partial d_2^*}{\partial b}$.

A.4 Derivation of the variance of $\hat{ASC}$

To solve $\frac{\partial ASC}{\partial a}$ we need to find $\frac{\partial AUC}{\partial a}$ and $\frac{\partial M}{\partial a}$. When $a > 0$ and $b > 0$, using the Leibniz integral, after some simplification we can show that

$$\frac{\partial AUC}{\partial a} = \left(\frac{1}{1-b} \right) \exp\left(\frac{a}{1-b} \right) \int_0^1 \frac{\left(\frac{x}{1-x} \right)^{\frac{1+b}{1-b}}}{\left[1 + \exp\left(\frac{a}{1-b} \right) \left(\frac{x}{1-x} \right)^{\frac{1+b}{1-b}} \right]^2} dx, \quad (A12)$$

and $\frac{\partial M}{\partial a}$ is equal to

$$\frac{\partial M}{\partial a} = \frac{\partial}{\partial a} \left[\frac{f(h(a,b), a, b)}{h(a,b)} \right] \frac{[h(a,b)]^2}{2} - \left(\frac{1}{1-b} \right) \exp\left(\frac{a}{1-b} \right) \int_0^{h(a,b)} \frac{\left(\frac{x}{1-x} \right)^{\frac{1+b}{1-b}}}{\left[1 + \exp\left(\frac{a}{1-b} \right) \left(\frac{x}{1-x} \right)^{\frac{1+b}{1-b}} \right]^2} dx \quad (A13)$$

where $h(a,b)$ is the solution of (A4) and

$$\frac{\partial}{\partial a} \left[\frac{f(h(a,b), a, b)}{h(a,b)} \right] = \frac{\frac{R}{1-b} \left[1 + \frac{1+b}{h(a,b)(1-h(a,b))} \frac{\partial h(a,b)}{\partial a} \right]}{(1+R)^2 h(a,b)} - \frac{R \left(\frac{\partial h(a,b)}{\partial a} \right)}{(1+R)[h(a,b)]^2},$$

$$\text{with } R = \exp\left(\frac{a}{1-b} \right) \left(\frac{h(a,b)}{1-h(a,b)} \right)^{\frac{1+b}{1-b}} \text{ and } \frac{\partial h(a,b)}{\partial a} = \frac{\exp\left(\frac{a}{1-b} \right) \left(\frac{h(a,b)}{1-h(a,b)} \right)^{\frac{1+b}{1-b}}}{\frac{1+b}{[1-h(a,b)]^2} \left[1 - \exp\left(\frac{a}{1-b} \right) \left(\frac{h(a,b)}{1-h(a,b)} \right)^{\frac{1+b}{1-b}} \right]}.$$

Similarly, to obtain $\frac{\partial ASC}{\partial b}$ we need to find $\frac{\partial AUC}{\partial b}$ and $\frac{\partial M}{\partial b}$. When $a > 0$ and $b > 0$, we have

$$\frac{\partial AUC}{\partial b} = \left(\frac{1}{1-b} \right)^2 \exp\left(\frac{a}{1-b} \right) \int_0^1 \frac{\left(\frac{x}{1-x} \right)^{\frac{1+b}{1-b}} \left[a + 2 \ln\left(\frac{x}{1-x} \right) \right]}{\left[1 + \exp\left(\frac{a}{1-b} \right) \left(\frac{x}{1-x} \right)^{\frac{1+b}{1-b}} \right]^2} dx \quad (A14)$$

and $\frac{\partial M}{\partial b}$ is equal to

$$\frac{\partial}{\partial b} \left[\frac{f(h(a,b), a, b)}{h(a,b)} \right] = \frac{[h(a,b)]^2}{2} - \left(\frac{1}{1-b} \right)^2 \exp\left(\frac{a}{1-b}\right) \int_0^{h(a,b)} \frac{\left(\frac{x}{1-x}\right)^{\frac{1+b}{1-b}} [a + 2 \ln(\frac{x}{1-x})]}{[1 + \exp(\frac{a}{1-b})(\frac{x}{1-x})^{\frac{1+b}{1-b}}]} dx \quad (\text{A15})$$

where

$$\frac{\partial}{\partial b} \left[\frac{f(h(a,b), a, b)}{h(a,b)} \right] = \frac{\frac{R}{(1-b)^2} [a + 2 \ln(\frac{h(a,b)}{1-h(a,b)}) + \frac{(1-b^2)}{h(a,b)(1-h(a,b))} \frac{\partial h(a,b)}{\partial b}]}{(1+S)^2 h(a,b)} - \frac{R(\frac{\partial h(a,b)}{\partial b})}{(1+R)[h(a,b)]^2}$$

$$\text{with } \frac{\partial h(a,b)}{\partial b} = \frac{\exp(\frac{a}{1-b})(\frac{h(a,b)}{1-h(a,b)})^{\frac{1+b}{1-b}} [a + 2 \ln(\frac{h(a,b)}{1-h(a,b)})] - \frac{2}{1-h(a,b)}}{\frac{(1-b^2)}{[1-h(a,b)]^2} [1 - \exp(\frac{a}{1-b})(\frac{h(a,b)}{1-h(a,b)})^{\frac{2b}{1-b}}]}.$$

The results when a and b take values in other ranges are derived similarly.

A.5 Derivation of the variance of $\hat{PLC}$ for homogeneous studies

From equation (8) we may derive that

$$\frac{\partial PLC_{\text{hom}}}{\partial a} = \pm \frac{\sqrt{2} \exp(\frac{a}{2})}{(1 + \exp(\frac{a}{2}))^2}. \quad (\text{A16})$$

Then the result follows.

A.6 Derivation of the variance of $\hat{ASC}$ for homogeneous studies

From equation (21) we may derive that

$$\frac{\partial AUC_{\text{hom}}}{\partial a} = \frac{\exp(a)}{(\exp(a) - 1)^3} [(\exp(a) + 1)a - 2(\exp(a) - 1)]. \quad (\text{A17})$$

Then the result follows.

REFERENCES

- Begg C. Meta-analysis methods for diagnostic accuracy. *Journal of Clinical Epidemiology* 2008; **61**:1081-1082.
- Erdreich LS, Lee ET. Use of relative operating characteristic analysis in epidemiology: a method for dealing with subjective judgment. *American Journal of Epidemiology* 1981; **114**:649-662.
- Franco E, Villa L, Rohan T, Ferenczy A, Petzl-Erler M, Matlashewski G. Design and methods of the Ludwig-McGill longitudinal study of the natural history of human papillomavirus infection and cervical neoplasia in Brazil Ludwig-McGill Study Group. *Rev Panam Salud Publica* 1999; **6**:223-233.
- Hanley JA. Receiver operating characteristic (ROC) curves. In *Encyclopedia of Biostatistics*, Armitage P, Colton T (eds). Wiley: Chichester, 1998; 3738-3745.
- Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; **143**:29-36.
- Hilden J. The area under the ROC curve and its competitors. *Medical Decision Making* 1991; **11**:95-101.
- Irwig L, Macaskill P, Glasziou P, Fahey M. Meta-analysis methods for diagnostic test accuracy. *Journal of Clinical Epidemiology* 1993; **48**:119-130.
- Lee WC, Hsiao CK. Alternative summary indices for the receiver operating characteristic curve. *Epidemiology* 1996; **7**:605-611.
- Macaskill P. Empirical Bayes estimates generated in a hierarchical summary ROC analysis agreed closely with those of a full Bayesian analysis. *Journal of Clinical Epidemiology* 2004; **57**:925-932.
- Midgette AS, Stukel TA, Littenberg B. A meta-analytic method for summarizing diagnostic test performances: receiver-operating-characteristic-summary point estimates. *Medical Decision Making* 1993; **13**:253-256.
- Moses LE, Shapiro DE, Littenberg B. Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. *Statistics in Medicine* 1993; **12**:1293-1316.
- Pepe M. *The Statistical Evaluation of Medical Tests for Classification and Prediction*. Oxford University Press: New York, 2004.
- Rutter CM, Gatsonis CA. Regression methods for meta-analysis of diagnostic test data. *Academic Radiology* 1995; **2**(Supplement 1):S48-S56.

- Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Statistics in Medicine* 2001; **20**:2865-2884.
- Scheidler J, Hricak H, Yu KK, Subak L, Segal MR. Radiological evaluation of lymph node metastases in patients with cervical cancer: a meta-analysis. *Journal of the American Medical Association* 1997; **278**:1096-1101.
- Seong HP, Jin MG, Chan HJ. Receiver operating characteristic (ROC) curve: practical review for radiologists. *Korean Journal of Radiology* 2004; **5**:11-18.
- Swets JA. Measuring the accuracy of diagnostic systems. *Science* 1988; **240**:1285-1293.
- Thompson ML, Zucchini W. On the statistical analysis of ROC curves. *Statistics in Medicine* 1989; **8**:1277-1290.
- Walter SD. Small sample estimation of log odds ratios from logistic regression and fourfold tables. *Statistics in Medicine* 1985; **4**:437-444.
- Walter SD. Properties of the summary receiver operating characteristic (SROC) curve for diagnostic test data. *Statistics in Medicine* 2002; **21**:1237-1256.
- Walter SD. The partial area under the summary ROC curve. *Statistics in Medicine* 2005; **24**:2025-2040.
- Walter SD, Sinuff T. Studies reporting ROC curves of diagnostic and prediction data can be incorporated into meta-analyses using corresponding odds ratios. *Journal of Clinical Epidemiology* 2007; **60**: 530-34.
- Zhou XH, Obuchowski NA, McClish DK. *Statistical Methods in Diagnostic Medicine*. Wiley: New York, 2002.