

Inference on the symmetry point-based optimal cut-off point and associated sensitivity and specificity with application to SARS-CoV-2 antibody data

A.M. Franco-Pereira^{1,2}, M.C. Pardo^{1,2}, C.T. Nakas^{3,4} and B. Reiser⁵

Abstract

In the presence of a continuous response test/biomarker, it is often necessary to identify a cut-off point value to aid binary classification between diseased and non-diseased subjects. The symmetry-point approach which maximizes simultaneously both types of correct classification is one way to determine an optimal cut-off point. In this article, we study methods for constructing confidence intervals independently for the symmetry point and its corresponding sensitivity, as well as respective joint nonparametric confidence regions. We illustrate using data on the generation of antibodies elicited two weeks post-injection after the second dose of the Pfizer/BioNTech vaccine in adult healthcare workers.

MSC: 62F10, 62G07, 65C05, 62P10.

Keywords: Empirical likelihood function, Empirical chi-square function, Box-Cox transformation, Confidence regions, Sensitivity, Specificity.

1. Introduction

Let X_1 and X_2 denote continuous response variables (biomarkers) for two user-defined groups (e.g. non-diseased versus diseased subjects), and let F_{X_1} and F_{X_2} be the cor-

¹ Department of Statistics and O.R., Complutense University of Madrid.

² Instituto de Matemática Interdisciplinar (IMI), Complutense University of Madrid, Plaza de Ciencias 3, 28040-Madrid, Spain.

³ Laboratory of Biometry, Department of Agriculture, Crop Production and Rural Environment, University of Thessaly, Phytokou street, 38446 Volos, Greece.

⁴ University Institute of Clinical Chemistry, Inselspital, Bern University Hospital, University of Bern, Freiburgstrasse, 3010 Bern, Switzerland.

Received: January 2023

Accepted: May 2023

responding probability distribution functions. Using a cut-off point c to decide that a subject is non-diseased when a marker measurement is less than c and that the subject is diseased otherwise, the specificity of the marker is $spec(c) = P(X_1 \leq c)$ and the sensitivity of the marker is $sens(c) = P(X_2 > c)$. We further make the standard assumption that larger values of the marker are more indicative of disease. The Receiver Operating Characteristic (ROC) curve is defined by $sens(c)$ versus $1 - spec(c)$ as c varies over the support of the response variable values. The symmetry point is the point c_s where $spec(c_s) = sens(c_s)$, which is where the ROC curve and the line $sens = 1 - spec$ intersect. The symmetry point approach for cut-off point selection has appeared in a wide range of recent applications in practice (see e.g. Arnone et al. 2020; Le, Ku and Jun, 2021; Sande et al. 2021; Sekgala et al. 2022) owing to its optimality properties rising from game theory considerations (Sanchez, 2017). However, in these studies, the estimated symmetry point is reported without confidence intervals, emphasizing the importance of having efficient and effective methods of statistical inference on the symmetry point cut-off and its sensitivity along with accessible software for implementation.

To our knowledge, the only two proposals for constructing confidence intervals (CIs) of the symmetry point-based optimal cutpoint and its associated sensitivity have been given by López-Ratón et al. (2016). They are based on the generalized pivotal quantity (GPQ) and the empirical likelihood (EL), respectively. The authors recommended the use of EL method when the distributions of healthy and diseased populations are unknown. Later, Adimari and Sinigaglia (2020) proposed a nonparametric method that provides joint confidence regions for the symmetry point-based optimal cutpoint and its associated sensitivity. Their method is also based on EL and uses the fact that the asymptotic distribution of the statistic they use has a chi-squared distribution with two degrees of freedom. We discuss an alternative to these nonparametric methods based on EL as well as parametric approaches for the construction of confidence intervals.

In the following section we present parametric and nonparametric approaches for the construction of confidence intervals for the symmetry point and its associated sensitivity, or equivalently its specificity, separately as well as methods for the construction of simultaneous confidence regions. Simulation studies comparing the different methods are presented in Section 3. In Section 4, an application to data of SARS-CoV-2 antibody levels is presented pertinent to the diagnosis of prior COVID-19 for possibly asymptomatic individuals. We end with a discussion.

2. Methods

2.1. Construction of confidence intervals: parametric approaches

Let $X_{11}, X_{12}, \dots, X_{1n_1}$ and $X_{21}, X_{22}, \dots, X_{2n_2}$ denote two random samples of sizes n_1 and n_2 taken from two independent normal distributions with mean μ_i and variance σ_i^2 , $i = 1, 2$, respectively. Under this assumption, it follows that the symmetry point c_s satisfies the following equation

$$\Phi\left(\frac{\mu_2 - \mu_1 + \sigma_1 \Phi^{-1}(x)}{\sigma_2}\right) = 1 - x$$

where $\text{spec}(c_s) = 1 - x$ and Φ is the cumulative distribution function of a variable following a standard normal distribution. After elemental algebra (see López-Ratón et al. (2016)), we obtain the following closed-form expression:

$$c_s = \frac{\mu_1 \sigma_2 + \mu_2 \sigma_1}{\sigma_1 + \sigma_2} \quad (1)$$

and

$$\text{spec}(c_s) = \text{sens}(c_s) = \Phi(\delta_s) = \Phi\left(\frac{\mu_2 - \mu_1}{\sigma_1 + \sigma_2}\right). \quad (2)$$

Both c_s and $\text{sens}(c_s)$ are estimated by substituting for the unknown $\mu_1, \mu_2, \sigma_1, \sigma_2$ in the above formulae their maximum likelihood estimates (MLEs), $\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}_1, \hat{\sigma}_2$. Sensitivity and specificity are proportions and thus they are bounded between zero and one. As a result, the normal approximation for the construction of confidence intervals described in the classical approach can be inadequate for small samples and may also result in intervals which exceed the bounds. To obtain a $(1-\alpha)\%$ confidence interval for $\text{sens}(c_s)$ we apply standard normal asymptotic theory on $\hat{\delta}_s$ which is not bounded, and use $\Phi\left(\hat{\delta}_s \pm z_{1-\alpha/2} \sqrt{\widehat{\text{Var}}(\hat{\delta}_s)}\right)$ where $z_{1-\alpha/2}$ refers to the $1 - \alpha/2$ percentile of the standard normal distribution. Since $\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}_1, \hat{\sigma}_2$ are all independent, using the delta method, we obtain

$$\widehat{\text{Var}}(\hat{c}_s) \approx \left(\frac{\partial \hat{c}_s}{\partial \mu_1}\right)^2 \widehat{\text{Var}}(\hat{\mu}_1) + \left(\frac{\partial \hat{c}_s}{\partial \mu_2}\right)^2 \widehat{\text{Var}}(\hat{\mu}_2) + \left(\frac{\partial \hat{c}_s}{\partial \sigma_1}\right)^2 \widehat{\text{Var}}(\hat{\sigma}_1) + \left(\frac{\partial \hat{c}_s}{\partial \sigma_2}\right)^2 \widehat{\text{Var}}(\hat{\sigma}_2)$$

where $\frac{\partial \hat{c}_s}{\partial \mu_1} = \frac{\hat{\sigma}_2}{\hat{\sigma}_1 + \hat{\sigma}_2}$, $\frac{\partial \hat{c}_s}{\partial \mu_2} = \frac{\hat{\sigma}_1}{\hat{\sigma}_1 + \hat{\sigma}_2}$, $\frac{\partial \hat{c}_s}{\partial \sigma_1} = \frac{\hat{\sigma}_2(\hat{\mu}_2 - \hat{\mu}_1)}{(\hat{\sigma}_1 + \hat{\sigma}_2)^2}$, $\frac{\partial \hat{c}_s}{\partial \sigma_2} = \frac{\hat{\sigma}_1(\hat{\mu}_1 - \hat{\mu}_2)}{(\hat{\sigma}_1 + \hat{\sigma}_2)^2}$ and

$$\widehat{\text{Var}}(\hat{\delta}_s) \approx \left(\frac{\partial \hat{\delta}_s}{\partial \mu_1}\right)^2 \widehat{\text{Var}}(\hat{\mu}_1) + \left(\frac{\partial \hat{\delta}_s}{\partial \mu_2}\right)^2 \widehat{\text{Var}}(\hat{\mu}_2) + \left(\frac{\partial \hat{\delta}_s}{\partial \sigma_1}\right)^2 \widehat{\text{Var}}(\hat{\sigma}_1) + \left(\frac{\partial \hat{\delta}_s}{\partial \sigma_2}\right)^2 \widehat{\text{Var}}(\hat{\sigma}_2)$$

where $\frac{\partial \hat{\delta}_s}{\partial \mu_1} = \frac{1}{\hat{\sigma}_1 + \hat{\sigma}_2}$, $\frac{\partial \hat{\delta}_s}{\partial \mu_2} = -\frac{1}{\hat{\sigma}_1 + \hat{\sigma}_2}$, $\frac{\partial \hat{\delta}_s}{\partial \sigma_1} = -\frac{\hat{\mu}_2 - \hat{\mu}_1}{(\hat{\sigma}_1 + \hat{\sigma}_2)^2}$, $\frac{\partial \hat{\delta}_s}{\partial \sigma_2} = -\frac{\hat{\mu}_2 - \hat{\mu}_1}{(\hat{\sigma}_1 + \hat{\sigma}_2)^2}$. The implied confidence interval for c_s is of the form $\hat{c}_s \pm z_{1-\alpha/2} \sqrt{\widehat{\text{Var}}(\hat{c}_s)}$. We refer to this approach as “ δ ”.

The assumption that the biomarkers are normally distributed can be quite restrictive leading to false results when it is significantly violated. A popular way of extending the parametric approach is the use of the Box-Cox transformation (Box and Cox, 1964) which has been previously employed in the ROC framework (e.g. Faraggi and Reiser

(2002); Fluss, Faraggi and Reiser (2005); Molodianovitch, Faraggi and Reiser (2006); Schisterman et al. (2008); Franco-Pereira, Nakas and Pardo (2020); López-Ratón et al. (2016)) and has been shown to perform very well. The Box-Cox transformation is defined by $X_i^{(\lambda)} = \frac{X_i^\lambda - 1}{\lambda}$ for $\lambda \neq 0$ and $X_i^{(0)} = \log(X_i)$ where it is assumed that $X_i^{(\lambda)} \sim N(\mu_i^{(\lambda)}, \sigma_i^{(\lambda)})$. The maximum likelihood estimate (MLE) $\hat{\lambda}$ of the common transformation parameter λ can be obtained by maximizing the profile likelihood function given in Franco-Pereira et al. (2021). We use $\hat{c}_s^{(BC)}$ and $\hat{\delta}_s^{(BC)}$ to denote the estimate of c_s and δ_s obtained above, but using the transformed observations. The estimator $\hat{c}_s^{(BC)}$ needs to be backtransformed to obtain an estimator of the symmetry point on the original scale. In this approach the “added” variation due to estimating the transformation is not taken into account following the rationale as Schisterman et al. (2008). We refer to this approach as “ δ -BC”.

Another possibility is to consider a bootstrap-based approach in order to obtain an estimate of the variance of \hat{c}_s and $\widehat{sens}(c_s)$ and thus to compute the $100(1 - \alpha)\%$ confidence interval for c_s and $sens(c_s)$ through the following steps:

Algorithm 1

1. Take a sample with replacement from X_1 and X_2 .
2. Carry out the Box-Cox transformation by maximizing the profile likelihood given in Franco-Pereira et al. (2021) for each bootstrap sample.
3. For $i = 1, 2$, calculate $\hat{\mu}_i^{(\lambda)}$ and $\hat{\sigma}_i^{(\lambda)}$, the MLE of μ_i and σ_i , respectively.
4. Calculate $\hat{c}_s^{(BC)}$ and $\hat{\delta}_s^{(BC)}$ in (1) and (2) by replacing the μ_i and σ_i with $\hat{\mu}_i^{(\lambda)}$ and $\hat{\sigma}_i^{(\lambda)}$.
5. Back-transform $\hat{c}_s^{(BC)}$ to obtain the current estimate for the symmetric point on the original scale, denoted by \hat{c}_s .
6. Repeat steps 1-5 B times. Then, based on the B values of \hat{c}_s and $\hat{\delta}_s^{(BC)}$, \hat{c}_{sb} and $\hat{\delta}_{sb}^{(BC)}$, derive the bootstrap estimate $\widehat{Var}_B(\hat{c}_s)$ and $\widehat{Var}_B(\hat{\delta}_s^{(BC)})$, respectively.
7. Construct the two-sided $100(1 - \alpha)\%$ confidence interval of c_s as $\hat{c}_s \pm z_{1-\alpha/2} \sqrt{\widehat{Var}_B(\hat{c}_s)}$ and $sens(c_s)$ as $\Phi\left(\hat{\delta}_s^{(BC)} \pm z_{1-\alpha/2} \sqrt{\widehat{Var}_B(\hat{\delta}_s^{(BC)})}\right)$.

We refer to this approach as “BC-AN”. The bootstrap estimates, \hat{c}_{sb} and $\hat{\delta}_{sb}^{(BC)}$ can also be used to obtain bootstrap percentile confidence intervals (“BC-PB”) as well as the bias corrected bootstrap confidence interval (“BC-bias”). Note again that these bootstrap estimators do take into account the variation due to the estimation of λ .

2.2. Construction of confidence intervals: nonparametric approaches

López-Ratón et al. (2016) constructed confidence intervals for the symmetry point and its associated sensitivity index using a parametric approach based on the generalized pivotal quantity (GPQ) and a non-parametric approach based on the empirical likelihood (EL). They recommended the use of the EL method when the distributions of healthy and diseased populations are unknown and this is the approach we are going to consider. Details of this procedure are presented therein. The method is based on the empirical likelihood function which is given by

$$l(sens(c_s), c_s) = 2n_1 \left\{ \widehat{F}_{X_1}(c_s) \log \frac{\widehat{F}_{X_1}(c_s)}{sens(c_s)} + \left(1 - \widehat{F}_{X_1}(c_s)\right) \log \frac{1 - \widehat{F}_{X_1}(c_s)}{1 - sens(c_s)} \right\} \quad (3)$$

$$+ 2n_2 \left\{ \widehat{F}_{X_2}(c_s) \log \frac{\widehat{F}_{X_2}(c_s)}{1 - sens(c_s)} + \left(1 - \widehat{F}_{X_2}(c_s)\right) \log \frac{1 - \widehat{F}_{X_2}(c_s)}{sens(c_s)} \right\}$$

where \widehat{F}_{X_i} is the Gaussian kernel estimate of the cumulative distribution function F_{X_i} , $i = 1, 2$, using the same bandwidth given in López-Ratón et al. (2016). We refer to this approach as “EL”. However, in many contexts, the chi-square test statistic works better than the likelihood ratio test statistic. Pardo, Lu and Franco-Pereira (2022) compared test statistics based on the empirical likelihood and chi-squared functions for testing monotone and umbrella orderings, and that based on the chi-square function was the most powerful. Pardo and Pardo (2008) found that the chi-square test statistic outperforms the classical loglikelihood test statistic for selecting a model from a sequence of Generalized Linear Models with binary data. Therefore, in this work we consider the procedure described by López-Ratón et al. (2016) substituting the empirical likelihood function with the empirical chi-square function

$$\Lambda(sens(c_s), c_s) = \frac{n_1 \left(\widehat{F}_{X_1}(c_s) - sens(c_s) \right)^2}{sens(c_s)(1 - sens(c_s))} + \frac{n_2 \left(\widehat{F}_{X_2}(c_s) - (1 - sens(c_s)) \right)^2}{sens(c_s)(1 - sens(c_s))} \quad (4)$$

in step 2 of their algorithm. A nonparametric estimator of $sens(c_s)$ is obtained which is in turn used to minimize Equation (4) in c_s and consider the minimum found as the nonparametric estimator of c_s . Then, we resample independently from the original pair of samples B times and repeating the above estimation procedure to obtain B bootstrap estimators of $sens(c_s)$ and c_s . Finally, these estimates are used to construct the CIs by the percentile method. We refer to this approach as “ECS”.

2.3. Construction of confidence regions

A joint region provides more precise information about the pair of parameters of interest $(sens(c_s), c_s)$ than the marginal confidence intervals do. In our simulation study reported below (Section 4) the parametric methods did not perform satisfactorily for the marginal

CIs obtained for c_s . As a result, we focus, in this section, on constructing nonparametric confidence regions for the symmetry point and its associated sensitivity index. Adimari and Sinigaglia (2020) proposed an approach based on the computation of the empirical distribution function from the data. As a consequence, the set

$$R_\alpha = \{(\text{sens}(c_s), c_s) : l(\text{sens}(c_s), c_s) \leq \chi_{2,\alpha}^2\}$$

with $l(\text{sens}(c_s), c_s)$ defined in (3) and $\chi_{2,\alpha}^2$ is such that $P(\chi_2^2 \geq \chi_{2,\alpha}^2) = \alpha$, is a confidence region with nominal coverage probability $1 - \alpha$ for the $(\text{sens}(c_s), c_s)$ point.

Since the asymptotic distribution of $\Lambda(\text{sens}(c_s), c_s)$ is the same as the asymptotic distribution of $l(\text{sens}(c_s), c_s)$ since $l(x, y) = \Lambda(x, y) + o_p(1)$, an alternative confidence region with nominal coverage probability $1 - \alpha$ is given by

$$R_\alpha^* = \{(\text{sens}(c_s), c_s) : \Lambda(\text{sens}(c_s), c_s) \leq \chi_{2,\alpha}^2\}$$

with $\Lambda(\text{sens}(c_s), c_s)$ defined in (4).

3. Simulation study

3.1. CIs

A large simulation study was conducted in order to compare the approaches for constructing confidence intervals described in Sections 2.1 and 2.2, namely: δ , δ -BC, BC-AN, BC-PB, BC-bias and ECS. We also compare our approaches with the one based on the empirical likelihood (EL) proposed by López-Ratón et al. (2016). We generated data from Normal, PowerNormal $(X_1^{-1/3}, X_2^{-1/3})$, LogNormal, Gamma and Mixed models. The parameters used for each of these scenarios are the same as those given in Table 1 of López-Ratón et al. (2016). We used sample sizes: $n_1 = n_2 = 30, 50, 100$ and the unequal sample size scenarios (20, 30), (50, 100) and (50, 300). The number of Monte Carlo replications utilized was $N=1000$ and $B=500$ for the bootstrap technique. Performances of each CI approach were assessed by coverage probability (CP) and mean of interval lengths (widths) for c_s and $\text{sens}(c_s)$, respectively.

Figures 1 and 3 summarize graphically the observed coverage probabilities and average widths of the CIs for each scenario for both c_s and sensitivity, respectively. In terms of coverage for c_s , except for the case of the mixture of normal distributions, we notice that the nonparametric methods perform substantially better than the parametric ones even in the normal case. Therefore, the width tends to be larger for the two nonparametric approaches. For normal, gamma and mixture models scenarios, both nonparametric methods, ECS and EL, are similar. ECS provides the best results for PowerNormal and EL for LogNormal. If we focus on the parametric approaches, the “BC-AN” approach outperforms the others except for the normal mixture scenario for which “ δ -BC” is the best but the spread of the observed average values is very large in comparison with the

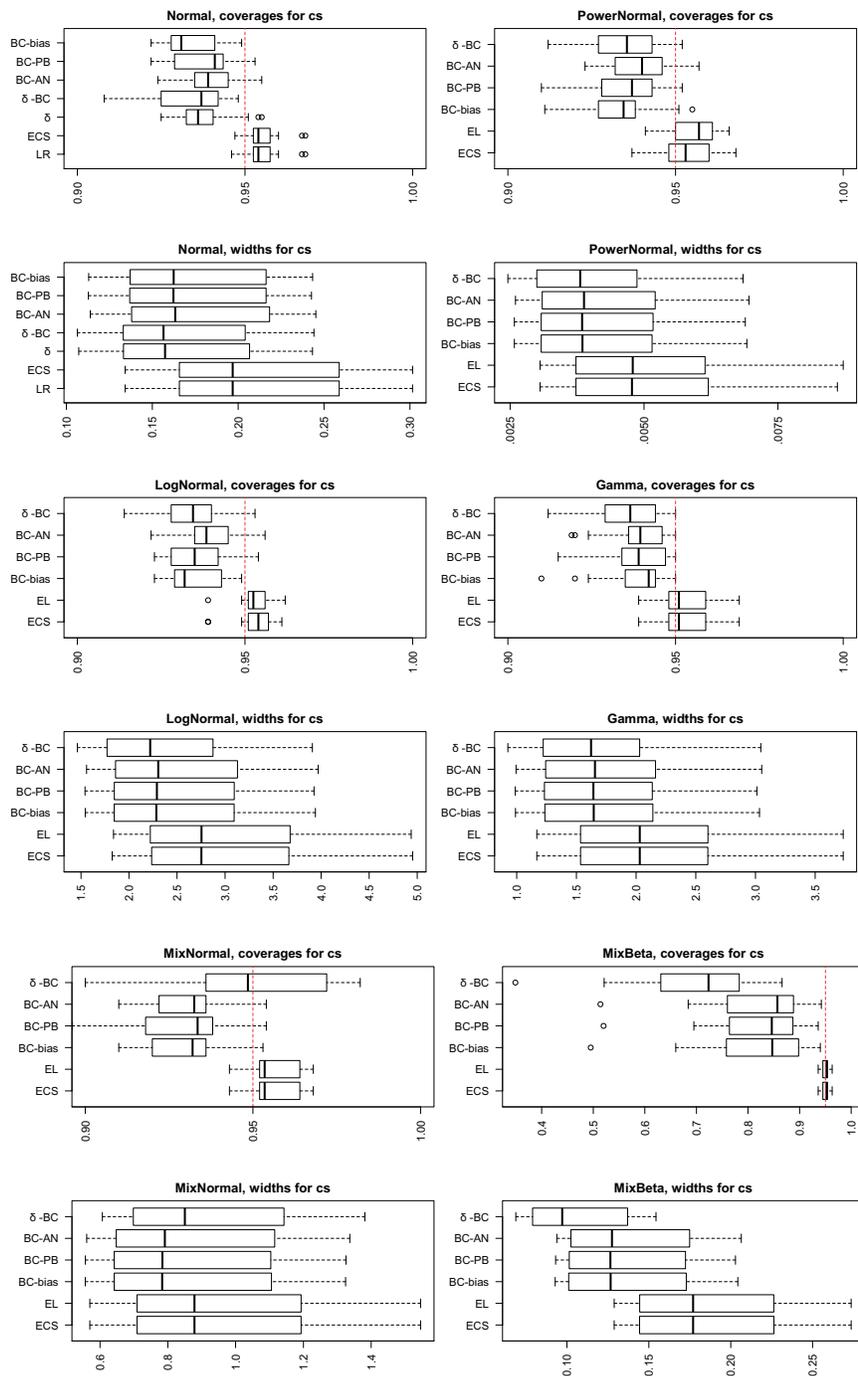


Figure 1. Boxplots of the coverages and average widths of the confidence intervals for c_s in the different scenarios considered in the simulation study.

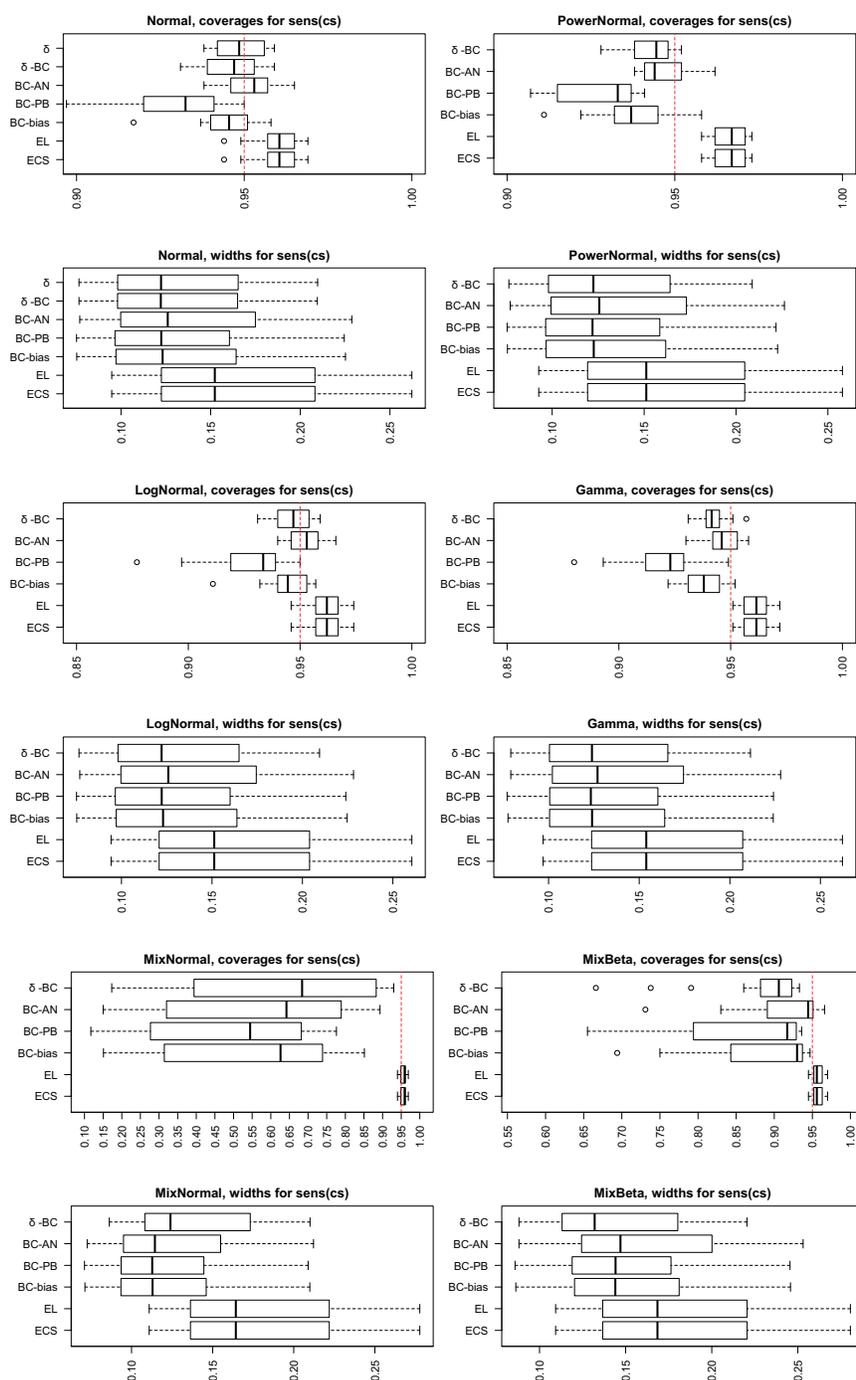


Figure 2. Boxplots of the coverages and average widths of the confidence intervals for $sens(c_s)$ in the different scenarios considered in the simulation study.

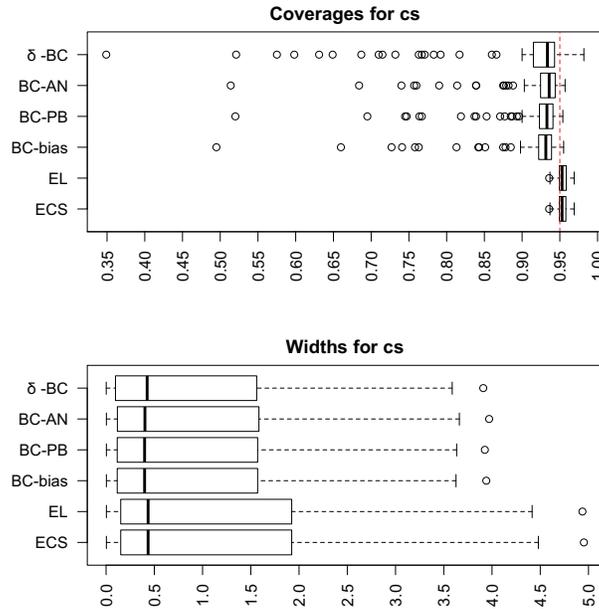


Figure 3. Boxplot of the coverages and average widths of the confidence intervals for c_s for all scenarios combined.

others. The parametric approaches tend to be somewhat conservative and their coverages tend to be more variable over the various scenarios and sample sizes sometimes exhibiting quite low coverages while the nonparametric methods have coverages which vary much less. However, looking at the coverages for $sens(c_s)$ except for the normal mixtures, BC-AN seems to do quite well having coverage closer to the nominal and shorter length than the nonparametric methods. In this case, the nonparametric methods tend to be more liberal having more than the nominal coverage which naturally leads to longer average widths.

Therefore, ECS and EL are recommended for constructing CIs for c_s and BC-AN for $sens(c_s)$ as can be seen in Figures 1 through 4, which provide a summary for all the methods merging all the scenarios into box-plots of the observed coverages and CI widths. It is important to take the coverage into account with a head-to-head comparison only being justified for similar coverages. However, the relationship between average length and coverage is not one-to-one, as a result, it is reported throughout for the sake of completeness in the Supplement (Tables 1-24).

3.2. CRs

A second simulation study was conducted to compare the approach based on the empirical likelihood function for constructing confidence regions proposed by Adimari and Sinagaglia (2020) with our proposal based on the chi-squared function given in (4). We generated data from Normal, LogNormal, Beta, Exponential, Gamma and Mixtures of normal distributions. The parameters used for each of these scenarios are the

same as those considered in Adimari and Sinagaglia (2020). We used sample sizes: $n_1 = n_2 = 20, 50, 100$ and the unequal sample size scenarios $(50, 20)$ and $(20, 50)$. The number of Monte Carlo replications was $N=10000$. The performance of each CR approach was assessed by the proportions of cases falling inside the confidence region, and mean of the confidence regions of the areas. In the simulation study we consider five different scenarios: two scenarios correspond to the normal model, for the third and fourth we use the beta and the gamma models, respectively. Finally, the fifth scenario corresponds to mixture models (see Table 25 in the Supplement).

The results of these simulations, for three levels of nominal coverage $1 - \alpha$, that is, 0.90, 0.95 and 0.99, are shown in Tables 26-41 in the Supplement. Tables 26-33 and Tables 34-41 give the estimated coverage probabilities and estimated areas of the confidence regions, respectively, obtained by using both methods presented in Section 2.3. for each scenario. As expected, the simulated coverage is closer to the nominal when the sample size increases. Our proposal generally provides results with observed coverage closer to the nominal level. However, the observed coverages of the confidence regions based on R_α^* are more variable than R_α for most of the scenarios. Figure 5 provides a graphical presentation of these results as well as a box-plot with the estimated coverage for all scenarios combined. In relation to the areas of the confidence regions, looking at Figure 6, it can be seen that they are very similar for both methods. Therefore, R_α^* approach produces confidence regions with a coverage closer to the nominal level without increasing their area.

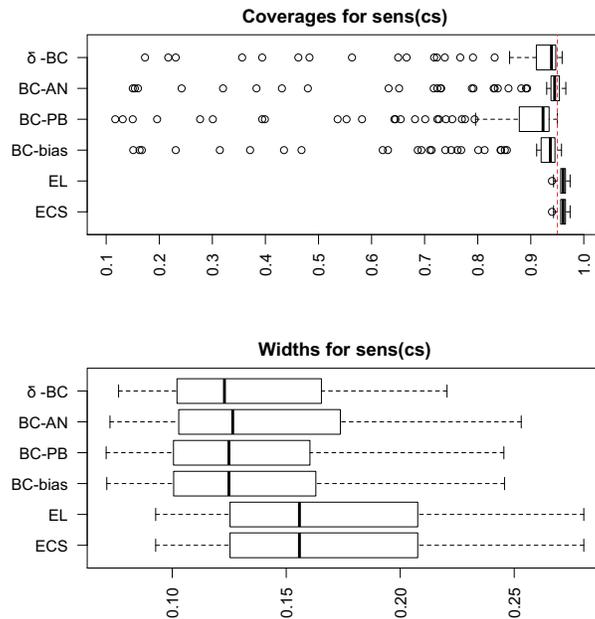


Figure 4. Boxplot of the coverages and average widths of the confidence intervals for $sens(c_s)$ for all scenarios combined.

4. Application to SARS-CoV-2 antibody data

We use part of the dataset measuring the generation of antibodies elicited two weeks post-injection after the second dose of the BNT162b2 mRNA vaccine in adult healthcare workers (Kontopoulou et al. 2021). All subjects were closely followed on a quasi-daily basis in order to flawlessly detect the presence of COVID-19. The estimation of a cut-off point that could be used in medical practice would aid diagnosis of prior COVID-19 infection, which in turn can guide decision making regarding correct classification of patients with symptoms pertinent to post COVID-19 syndrome (long covid). The data consist of 289 subjects without prior COVID-19 and 50 subjects with confirmed prior COVID-19. Antibody data do not conform to normality assumptions in general (Anderson and Darling, 1952) normality test p -value $< 2.2e - 16$ for controls and 0.002188 for cases) and a log10 transformation is a typical remedy in order to pursue formal hypothesis testing (Horne-Dale, 1995). The Box-Cox transformation provides a straightforward approach for an optimal estimation of the power transformation to normality. Kernel density estimators illustrating measurements before and after the Box-Cox transformation ($\hat{\lambda} = 0.235$) are given in Figure 7 (a) and (b), respectively. In addition, we have included the corresponding estimation of the densities under the binormal model (after the Box-Cox transformation) and its corresponding ROC curve, whose associated AUC value is 0.716, in Figure 7 (c) and (d), respectively.

Results regarding the cut-off point based on the symmetry point methods presented in Section 2 are given in Table 1, suggesting that antibody measurements above around 22300 suggest prior COVID-19 infection, when antibodies are measured two weeks after the second dose of the BNT162b2 mRNA vaccine, with sensitivity around 65%. The classification of patients, with symptoms pertinent to post COVID-19/long Covid syndrome is to be taken with caution in practice. Antibody levels two weeks post-injection after the second dose of the BNT162b2 mRNA vaccine is a significant marker but cannot be used as a standalone test in practice given its very limited utility nowadays and its moderate performance. Corresponding confidence regions (CRs) and their illustration are given in Table 2 and Figure 8. Confidence regions are somewhat tighter using the R_{α}^* approach, providing an apparent higher estimation accuracy.

Table 1. Estimates and 95% CIs for the different methods.

Method	\hat{c}_s	95% CI (\hat{c}_s)	$\widehat{sens}(c_s) = \widehat{spec}(c_s)$	95% CI $sens(\hat{c}_s)$
δ	25234.90	(22814.74, 27655.06)	0.641	(0.578, 0.703)
δ -BC	22304.13	(20103.32, 24504.93)	0.656	(0.601, 0.711)
BC-AN	22220.90	(19888.72, 24553.08)	0.656	(0.597, 0.714)
BC-PB	22299.64	(20103.53, 24495.76)	0.661	(0.608, 0.713)
BC-Bias	22251.90	(20119.51, 24384.29)	0.651	(0.599, 0.704)
EL	22296.64	(19597.31, 24995.98)	0.655	(0.591, 0.719)
ECS	22296.64	(19597.31, 24995.98)	0.655	(0.591, 0.719)

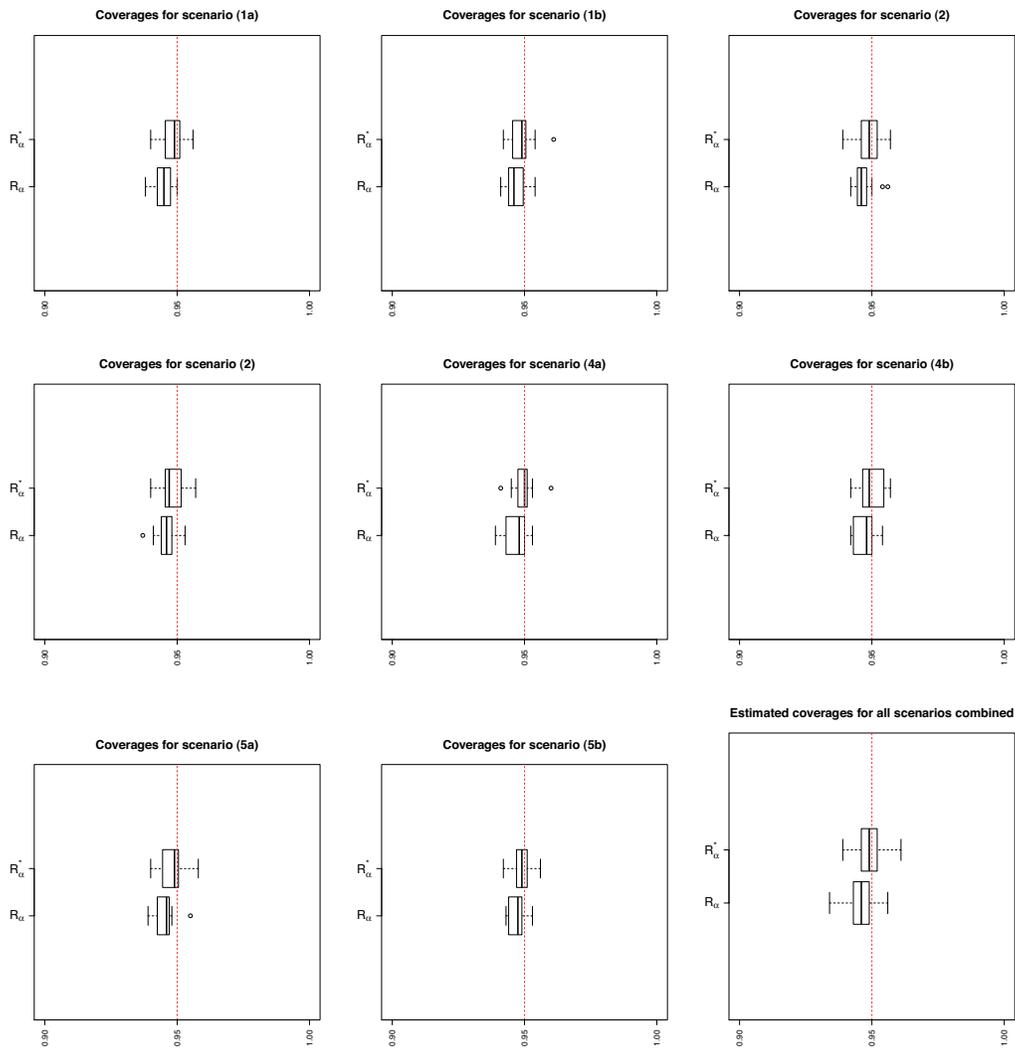


Figure 5. Boxplots of the estimated coverages of the confidence regions at 95% confidence level for all scenarios considered in the simulation study, including a boxplot (the last one) with the results of all scenarios combined.

Table 2. Estimates of 90,95,99% CRs for the different methods.

Area	90% CR	95% CR	99% CR
R_α	463.62	621.10	946.14
R_α^*	459.47	603.13	939.89

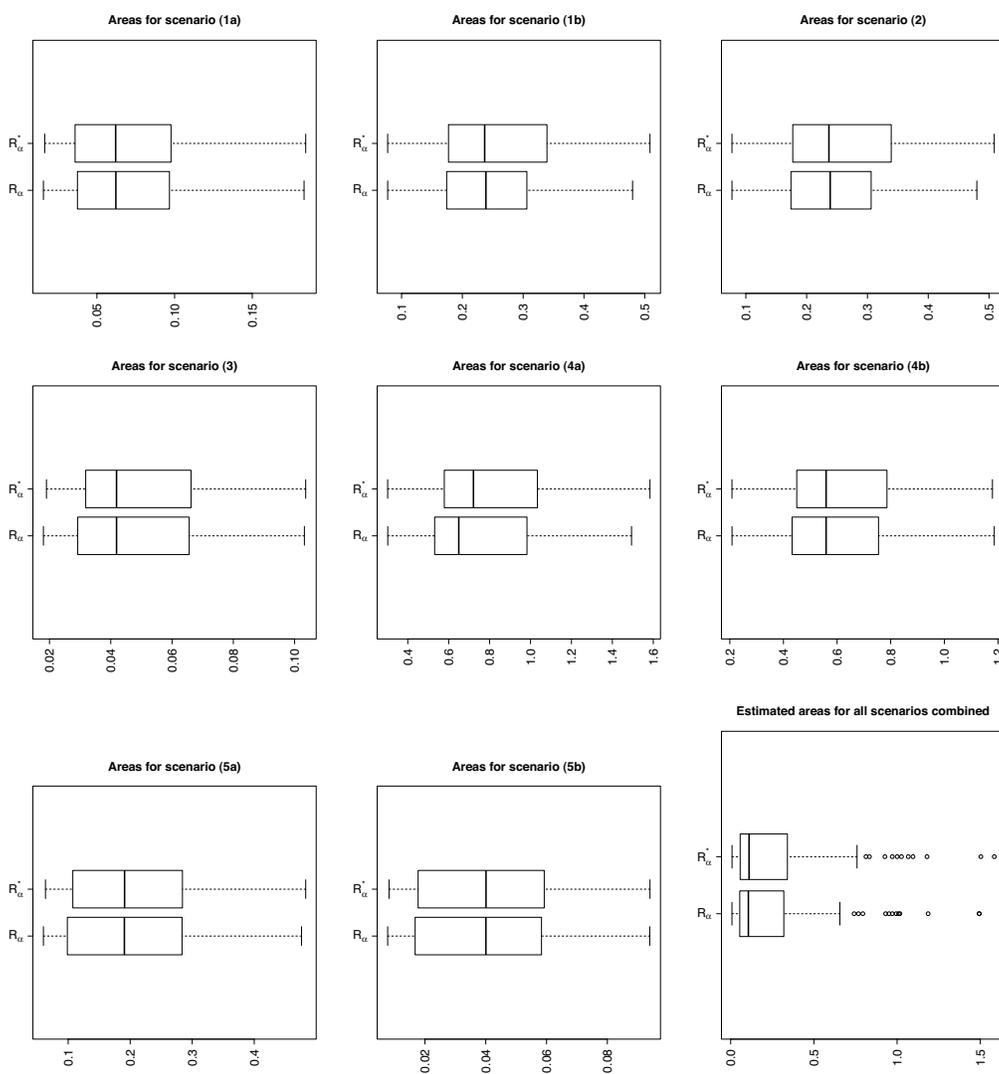


Figure 6. Boxplot of the estimated areas of the confidence regions at 95% confidence level for all scenarios considered in the Simulation Study, including a boxplot (the last one) with the results of all scenarios combined.

5. Discussion

When considering a continuous biomarker/score clinicians are in need of a cut-off point/optimal threshold to classify subjects into one of the two diagnostic groups under consideration. One such method is based on the symmetry point. This approach has theoretical

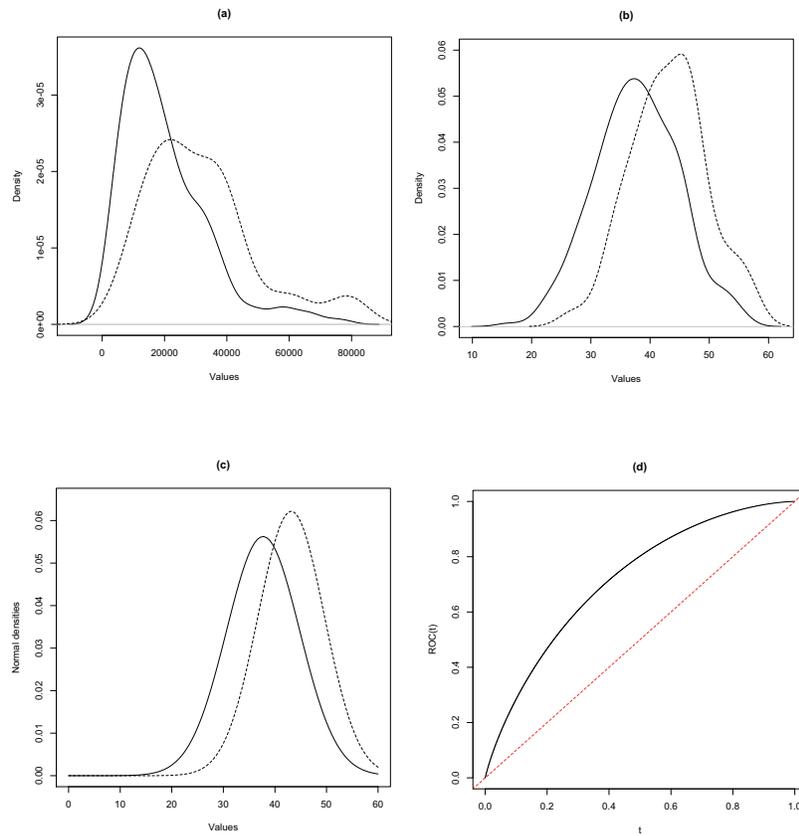


Figure 7. Top row: Kernel density estimators before (a) and after (b) BC for the SARS-CoV-2 antibody data. Bottom row: Estimated densities (c) and corresponding ROC curve (d) under the binormal model, after BC.

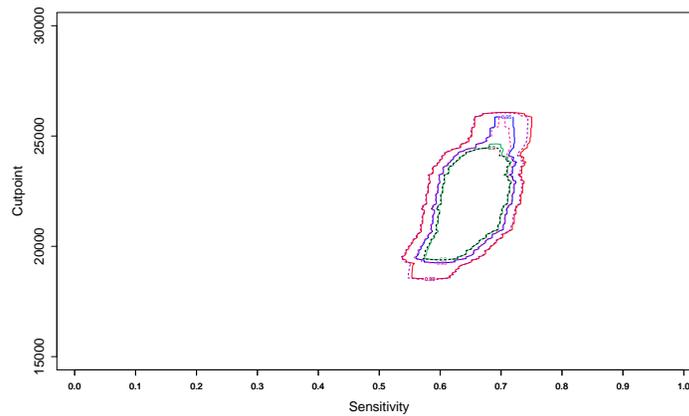


Figure 8. CR contours for the SARS-CoV-2 antibody data. Solid lines correspond to R_α and dashed ones to R_α^* . Colours red and purple are used for the 99% CRs; blue and pink for the 95% CRs and green and black for the 90% CRs.

support from a game theory minimax perspective and has important robustness features (Sanchez, 2017). The symmetry point has been used in applications but typically without the addition of confidence intervals. In this work, we have considered both marginal CIs for the symmetry point and its associated sensitivity and CRs for these jointly. For the marginal CIs we examined both parametric and nonparametric approaches. The parametric approaches are based on the binormal model and implemented using the Box-Cox transformation for cases when normality assumptions are not met. The BC-AN method worked quite well for the associated sensitivity for a wide range of distributions except for the mixtures. Two nonparametric methods were examined. The first due to López-Ratón et al. (2016) is based on empirical likelihood while the second is based on the chi-square test statistic and has been found to work well in other contexts. Both of these give very similar results and work quite well both for the symmetry point and its corresponding sensitivity although somewhat conservative for most scenarios. Due to the poor performance of the parametric CIs for c_s we only examined nonparametric confidence regions for the symmetry point and its associated sensitivity index. The first due to Adimari and Sinigaglia (2020) is based on empirical likelihood. We proposed an alternative again based on a chi-square test statistic. Our simulations indicate that although these two methods perform similarly our proposal generally provides results with observed coverage closer to the nominal level. The availability of CIs and CRs for the symmetry point approach should help practitioners using this method in their data analyses. Software for carrying out these procedures is available from the first author. We illustrated these procedures using part of the dataset from a published study on SARS-CoV-2 antibody levels post vaccination.

In addition to the symmetry point approach there are many other methods proposed in the literature for obtaining the optimal cut-off point value. Commonly seen methods include maximizing the Youden Index (Bantis, Nakas and Reiser, 2019) or its weighted version (Schisterman et al., 2005), point closest to the (0,1) corner (Perkins and Schisterman, 2006) and maximizing the product of sensitivity and specificity (Liu, 2012) among others. With many possible methods there is an inherent problem in choosing the appropriate method for the selection of an optimal cutoff point that can be used in everyday practice. Weights reflecting the relative importance of sensitivity and specificity can be introduced in a cost-benefit tradeoff approach. Researchers can be expected to potentially run a large number of cut-off point selection approaches and estimate cut-off points along with corresponding CIs prior to final decision making. It is difficult to say which approach is more clinically relevant. No automatic procedure for such a choice currently exists and it is not clear if this is possible at all. Future research on this issue would be beneficial for practical applications in clinical problems. An initial step in this direction would be a detailed comparison of cut-off point selection methods along with the assessment of the robustness of corresponding CIs.

Acknowledgments

This work was supported by grants PID2019-104681RB-I00. Data courtesy of Dr Konstantina Kontopoulou.

References

- Anderson, T.W. and Darling, D.A. (1952). Asymptotic Theory of Certain “Goodness of Fit” Criteria Based on Stochastic Processes. *Ann. Math. Statist.*, 23, 193-212.
- Arnone, E., Cucchi M., Dal Gesso, S., Petitta, M. and Calmanti, S. (2020). Droughts Prediction: a Methodology Based on Climate Seasonal Forecasts. *Water Resources Management*, 34, 4313-4328.
- Adimari, G. and Sinigaglia, A. (2020). Nonparametric confidence regions for the symmetry point-based optimal cutpoint and associated sensitivity of a continuous-scale diagnostic test. *Biometrical Journal*, 62, 1463-1475.
- Bantis, L.E., Nakas, C.T. and Reiser, B. (2019). Construction of confidence intervals for the maximum of the Youden index and the corresponding cutoff point of a continuous biomarker. *Biometrical Journal*, 61(1), 138-156.
- Box, G.E. and Cox, D.R. (1964). An analysis of transformations. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 26, 211-252.
- Faraggi, D. and Reiser, B. (2002). Estimation of the area under the ROC curve. *Statistics in Medicine*, 21, 3093-3106.
- Fluss, R., Faraggi, D. and Reiser, B. (2005). Estimation of the Youden Index and its Associated Cutoff Point, *Biometrical Journal*, 47, 458-472.
- Franco-Pereira, A.M., Nakas, C.T. and Pardo, M.C. (2020). Biomarker assessment in ROC curve analysis using the length of the curve as an index of diagnostic accuracy: the binormal model framework. *AStA Advances in Statistical Analysis*, 104, 625-647.
- Franco-Pereira, A.M., Nakas, C.T., B. Reiser and Pardo, M.C. (2021). Inference on the Overlap Coefficient: The binormal approach and alternatives. *Statistical Methods in Medical Research*, 30, 2672-2684.
- Horne-Dale, A. (1995). The Statistical Analysis of Immunogenicity Data in Vaccine trials. *Annals of the New York Academy of Sciences*, 754, 329-346.
- Kontopoulou, K., Ainatzoglou, A., Nakas, C.T., Ifantidou, A., Goudi, G., Antoniadou, E., Adamopoulos, V., Papadopoulos, N., and Papazisis, G. (2021). Second dose of the BNT162b2 mRNA vaccine: Value of timely administration but questionable necessity among the seropositive. *Vaccine*, 39, 5078-5081.
- Le, R., Ku, H. and Jun, D. (2021). Sequence-based clustering applied to long-term credit risk assessment. *Expert Systems With Applications*, 165, 113940.
- Liu, X. (2012). Classification accuracy and cut point selection. *Statistics in Medicine*, 31(23), 2676-2686.

- López-Ratón, M., Cadarso-Suárez, C., Molanes-López, E.M. and Letón, E. (2016). Confidence intervals for the symmetry point: An optimal cutpoint in continuous diagnostic tests. *Pharmaceutical Statistics*, 15, 178-192.
- Molodianovitch, K., Faraggi, D. and Reiser, B. (2006). Comparing the areas under two correlated ROC curves: parametric and non-parametric approaches. *Biometrical Journal*, 48, 745-757.
- Pardo, J.A. and Pardo, M.C. (2008). Minimum phi-divergence estimator and phi-divergence statistics in Generalized linear models with binary data. *Methodology and Computing in Applied Probability*, 10, 357-379.
- Pardo, M.C., Lu, Y. and Franco-Pereira, A.M. (2022). Extensions of empirical likelihood and chi-squared-based tests for ordered alternatives. *Journal of Applied Statistics*, 49, 24-43.
- Perkins, N.J. and Schisterman, E.F. (2006). The inconsistency of “optimal” cutpoints obtained using two criteria based on the receiver operating characteristic curve. *American Journal of Epidemiology*, 163(7), 670-675.
- Sanchez, IE. (2017) Optimal threshold estimation for binary classifiers using game theory [version 3; peer review: 3 approved] *F1000Research 2017*, 5(ISCB Comm J): 2762.
- Sande, S.Z., Seng, L., Li, J. and D' Agostino, R. (2021). Statistical Learning in Medical Research with Decision Threshold and Accuracy Evaluation. *Journal of Data Science*, 19, 634-657.
- Schisterman, E.F., Faraggi, D., Reiser, B. and Hu, J. (2008). Youden Index and the optimal threshold for markers with mass at zero. *Statistics in medicine*, 27, 297-315.
- Schisterman, E.F., Perkins, N.J., Liu, A. and Bondell, H. (2005). Optimal cut-point and its corresponding Youden Index to discriminate individuals using pooled blood samples. *Epidemiology*, 73-81.
- Sekgala, M.D., Opperman, M., Mpahleni, B. and Mchiza, Z.J.-R. (2022). Anthropometric indices and cut-off points for screening of metabolic syndrome among South African taxi drivers. *Front. Nutr.* 9:974749.

