# Adjusting ROC Curve for Covariates with AROC R Package

Francisco Machado e Costa[1(✉)] and Ana Cristina Braga[2]

[1] School of Engineering, University of Minho, Campus de Gualtar, Braga, Portugal
frmachadoecosta@gmail.com
[2] ALGORITMI Centre, University of Minho, Campus de Azurém,
Guimarães, Portugal
acb@dps.uminho.pt

**Abstract.** The ability of a medical test to differentiate between diseased and non-diseased states is of vital importance and must be screened by statistical analysis for reliability and improvement. The receiver operating characteristic (ROC) curve remains a popular method of marker analysis, disease screening and diagnosis. Covariates in this field related to the subject's characteristics are incorporated in the analysis to avoid bias. The covariate adjusted ROC (AROC) curve was proposed as a method of incorporation. The AROC R-package was recently released and brings various methods of estimation based on multiple authors work.

The aim of this study was to explore the AROC package functionality and usability using real data noting its possible limitations. The main methods of the package were capable of incorporating different and multiple variables, both categorical and continuous, in the AROC curve estimation. When tested for the same data, AROC curves are generated with no statistical differences, regardless of method.

The package offers a variety of methods to estimate the AROC curve complemented with predictive checks and pooled ROC estimation. The package offers a way to conduct a more thorough ROC and AROC analysis, making it available for any R user.

**Keywords:** Receiver operator characteristic curve · Covariate adjustment · Diagnostic test · Biostatistics · Software tool

## 1 Introduction

Medical tests meant for use in routine practice must be rigorously scrutinized for their accuracy with statistical analysis before approval.

Differentiating between healthy and diseased populations or, more broadly, control and cases is the fundamental property of any test and of fundamental importance in clinical practice. The receiver operating characteristic (ROC) curve is the most popular statistical tool when dichotomizing test results, particularly with a degree of subjectivity, such as medical imaging [11] and is used

broadly in biomedical informatics research to evaluate classification and prediction models for decision support, diagnosis, and prognosis [8].

The ROC curve is defined as a plot of True Positive Fraction ($TPF$), or sensitivity, and False Positive Fraction ($FPF$), 1- specificity, pairs obtained by varying threshold $c$. Defining $Y_D$ and $Y_{\bar{D}}$ as continuous variables for diseased and non-diseased groups respectively with cumulative distribution functions $F_D$ and $F_{\bar{D}}$, we assume all test outcomes greater than $c$ belong to the diseased group, with $c \in \mathbb{R}$. Subsequently, each given $c$ will determine TPF, $TPF(c) = Pr(Y_D \geq c) = 1 - F_D(c)$ and similarly FPF, $FPF(c) = Pr(Y_{\bar{D}} \geq c) = 1 - F_{\bar{D}}(c)$.

The ROC curve is then defined as all TPF-FPF pairs, $ROC(\cdot) = \{(FPF(c), TPF(c)), c \in \mathbb{R}\}$ [13]. By converting $FPF$ at threshold $c$ to $t$, such as, $t = FPF(c) = 1 - F_{\bar{D}}(c)$, the ROC curve is defined as $\{(t, ROC(t)) : t \in [0,1]\}$ [16], where

$$ROC(t) = Pr\{Y_D > F_{\bar{D}}^{-1}(1-t)\} = 1 - F_D\{F_{\bar{D}}^{-1}(1-t)\}, \quad 0 \leq t \leq 1 \quad (1)$$

Clinical risk index for babies (CRIB) is a risk assessment tool used in neonatal intensive care units (NICUs) for infants born with less than 31 week gestation or 1500 g and lower birth weight [10,12]. Along with the Score for Neonatal Acute Physiology (SNAP), these scoring systems and their updates have served as prediction tools more accurate to the previous weight or gestational age univariate predictors.

CRIB score uses six different variables obtained routinely during the first 12 hours of life, namely, birthweight, gestational age, the presence of congenital malformation (excluding inevitably lethal congenital malformations) and indices of physiological status [1] resulting in a score between 0 and 24.

Matching sensitivity and specificity, defining an acceptable trade-off between the two to better characterize test results is, while not its only application, the main focus of ROC curve studies, and CRIB scores have been explored in this regard [1,2]. However in any diagnostics test a number of factors such as the disease characteristics and specimen features, might affect the marker and its accuracy. In statistics these unaccounted variables are labeled covariates. Incorporating covariate information in ROC analysis is imperative, otherwise biased or oversimplified conclusions might be made about the tests accuracy [7].

In recent years several different mathematical approaches have been introduced to better equate covariates in ROC curve methodology.

By denoting $\mathbf{X}_D$ and $\mathbf{X}_{\bar{D}}$ as diseased and non-diseased vector of covariates of interest and a covariate value $\mathbf{x}$, the covariate-specific ROC curve is built as

$$ROC(t|x) = Pr\{Y_D > F_{\bar{D}}^{-1}(1-t|\mathbf{X}_{\bar{D}} = \mathbf{x})|\mathbf{X}_D = \mathbf{x}\} \quad (2)$$

$$1 - F_D\{F_{\bar{D}}^{-1}(1-t|\mathbf{X}_{\bar{D}} = \mathbf{x})|\mathbf{X}_D = \mathbf{x}\} \quad (3)$$

The covariate specific ROC curve builds different curves and displays a different test accuracy for each value of $\mathbf{x}$ that, while helpful in understanding optimal and sub optimal populations for the test, lacks a global measurement

for the covariate effect. The covariate adjusted ROC curve, AROC, was proposed to mend this issue [7], defined as,

$$AROC(t) = \int ROC(t|\mathbf{x})dH_D(\mathbf{x}) \tag{4}$$

where,

$$H_D(x) = Pr(\mathbf{X}_D \leq \mathbf{x}),$$

the cumulative distribution function of $\mathbf{X}_D$. Therefore the AROC is a weighted average of covariate specific ROC curves according to the distribution of the covariates in the diseased group [16]. Janes and Pepe [7] also demonstrated that the AROC curve can also be expressed as

$$AROC(t) = Pr\{Y_D > F_{\bar{D}}^{-1}(1 - t|\mathbf{X}_D)\} = Pr\{1 - F_{\bar{D}}(Y_D|\mathbf{X}_D)) \leq t\}, \tag{5}$$

proving that the AROC summarizes the covariate-specific performance of the test.

The AROC R-package developed by Rodriguez-Alvarez and Inacio de Carvalho [16] implements different methods of computing covariate information in ROC curve construction and two additional methods of calculating marginal/polled ROC curve. This package has additional value considering the lag on tool development on this sub-field.

This article will explore these different methods on a theoretical and practical basis using CRIB scores of a neonatal database from the Portuguese National Registry on low weight newborns between 2010 and 2012.

## 2    Methodology

The aim of this study is to explore the AROC package functionality and usability using real data noting its possible limitations and generating preliminary findings on CRIB data that will be explored in the future.

### 2.1    Dataset

As mentioned the dataset used is part of the Portuguese National Registry on low weight newborns between 2010 and 2012 and was made available for research purposes. The original data included possible confounding observations such as repeated id's and twins that were removed to ensure no unaccounted for variables.

After ensuring all id's were unique these were promptly removed along with any possible identifiable features to abide by EU anonymity and data protection standards. A random sample of 50% was taken resulting in a dataset with 1093 unique observations and 17 columns including gestational time in weeks, the mothers age, biological sex of the infant, CRIB score, survival and other possibly relevant covariates and used for the remainder of the study.

For replication purposes all relevant data was shared to Mendeley Data and is available at data.mendeley.com/datasets/jsmgcmfmdx [9].

## 2.2   AROC Package Method Exploration

The AROC package for covariate adjusted Receiver Operating Characteristic curve inference offers a comprehensive guide of use present in the CRAN repository. In the remainder of this section we will replicate the examples provided in the documentation with the CRIB dataset, compare the different methods and explore some of the theoretical details of each method.

We begin by importing the dataset, setting a seed for replication purposes and constructing the specnaomit() function to remove NAs on columns under analysis, this is a necessary step since the methods are not built to handle NA values.

```
library(AROC)
set.seed(0824)
neonatal <- read.csv('NeonatalPortugal.csv',
                      sep = ';')

specnaomit <- function(col1, col2, df) {
  newdf <- df

  if (any(is.na(newdf[[col1]]))) {
    newdf <- newdf[!is.na(newdf[[col1]]),]

  }
  if (any(is.na(newdf[[col2]]))) {
    newdf <- newdf[!is.na(newdf[[col2]]),]

  }

  newdf

}

# Crib with Sex as covariate
neonatalCribSex <- specnaomit('sex',
                              'crib',
                              neonatal)

# Crib with Mother's age as covariate
neonatalCribAM <- specnaomit('age_mother',
                             'crib',
                             neonatal)
```

**Nonparametric Bayesian.** The *AROC.bnp* function estimates the AROC curve using the Bayesian nonparametric method. Working with Eq. 5 we say,

$$AROC(t) = Pr\{Y_D > F_{\bar{D}}^{-1}(1 - t|\mathbf{X}_D)\}$$
$$= Pr\{1 - F_{\bar{D}}(Y_D|\mathbf{X}_D)) \le t\}$$
$$= Pr(U_D \le t), \quad 0 \le t \le 1, \tag{6}$$

where $U_D = 1 - F_{\bar{D}}(Y_D|\mathbf{X}_D)$, a placement value of the test outcome in the diseased population or the standardization of $Y_D$ to the conditional distribution of $Y_{\bar{D}}$ making the AROC a cumulative distribution function of $U_D$. This method first models the conditional distribution of test outcomes in the nondiseased group, $F_{\bar{D}}$ using a B-splines dependent Dirichlet process mixture of Normals model followed by modeling $U_D$ and it's cumulative distribution using a non parametric regression model through Bayesian bootstrap [16].

For this demonstration we attempt to see the relation between the CRIB score and the sex of the baby, a categorical variable, that could indicate a possible bias on estimating survivability based on sex in the score system. To use this method the only mandatory arguments are *formula.healthy*, an R formula object to define the relation between variables, in this case the 'crib' and 'sex', *group*, that defines the result column, in this case if the infant survived, *tag.healthy*, what identifies the healthy i.e. survival population in the group/result column and *data*, the dataset being used. Further customization is possible with the Bayesian parameters however this will not be explored, the only extra arguments were meant to reduce the default run time.

```
#------ bnp method ------#

bnpCribSex <-
  AROC.bnp(formula.healthy ='crib ~ sex',
                    group = 'status',
                    tag.healthy = 0,
                    data = neonatalCribSex,
                    nsim = 5000,
                    nburn = 1000)

summary(bnpCribSex)

plot(bnpCribSex)
```

The plot functionality allows us to see the computed curve (Fig. 2A) that follow a simplistic visual however the saved object preserves the curve points that allow for any plot customization R programmers use.

**Semiparametric Bayesian.** The *AROC.bsp* method is very similar to the previous, non parametric method, in both construction and theory, where this models $F_{\bar{D}}$ using a normal linear regression, making it a counterpart to Janes and Pepe [7] frequentist model and its AROC package method - AROC.sp [14].

```
#------ bsp method ------#

bspCribSex <-
  AROC.bsp(formula.healthy ='crib ~ sex',
           group = 'status',
           tag.healthy = 0,
           data = neonatalCribSex)

summary(bspCribSex)

plot(bspCribSex)
```

The object generated by this method can, again, be plotted (Fig. 2B) similarly to the previous method

**Semiparametric Frequentist.** This semiparametric frequencist method arguments and construction are fairly similar to the previous methods but, as previously hinted, uses a semiparametric location regression model for $Y_{\bar{D}}$ to estimate $F_{\bar{D}}$ and estimates outer probability empirically [6,14] making it less computationally heavy, i.e. faster in comparison. Another advantage is being able to provide direct insight on covariate influence over the test/marker with the fit model's parameters to the crib and sex example we've seen previously.

```
spCribSex <-
  AROC.sp(formula.healthy ='crib ~ sex',
          group = 'status',
          tag.healthy = 0,
          data = neonatalCribSex)

summary(spCribSex)
summary(spCribSex$fit.h)

plot(spCribSex)
```

In this instance, calling the *summary()* function on the *fit.h* element give us an extensive look at test statistics.

**Nonparametric Kernel.** The *AROC.kernel*, an earlier proposed model by Rodriguez-Alvarez *et al.* [15] is, as the name implies, a kernel based method. Test outcomes for the non-diseased group are modeled with a location-scale regression model where both regression and variance functions are estimated using Nadaraa-Watson local estimators that in turn are used to compute standardised residuals to model $U_D$ and estimate the AROC curve, $\widehat{AROC}(t)$ [14–16].

The package author notes that, for now, this method, unlike the previous, can only handle a single continuous covariate, for this reason the sex variable is unsuitable for testing and is replaced by the age of the mother, *age_mother* to analyze this covariates influence in the CRIB score system and its outcome. Additionally *AROC.sp* is used to offer additional information.

```
#------ kernel method ------#

kernelCribAM <-
  AROC.kernel(marker = 'crib',
              covariate = 'age_mother',
              group = 'status', tag.healthy = 0,
              data = neonatalCribAM,
              B = 10) #time

summary(kernelCribAM)

plot(kernelCribAM)

#------ sp method ------#

spCribAM <-
  AROC.sp(formula.healthy ='crib ~ age_mother',
          group = 'status',
          tag.healthy = 0,
          data = neonatalCribAM)

summary(spCribAM)
summary(spCribAM$fit.h)

plot(spCribAM)
```

Note that for the *AROC.kernel*, not just the default arguments are different from the remaining methods but also that it was necessary to shorten the bootstrap resample process (B) from its default 1000 value, this is due to the execution time of this method being the longest of all the AROC-package functions.

**Posterior Predictive Checks (PPC).** All previous methods are meant to construct and analyze the AROC curve and the behavior of the incorporated covariates, for the remainder of this section the methods focus on posterior predictive checks and pooled ROC estimation.

Both *predictive.checks.AROC.bnp* and *predictive.checks.AROC.bsp* are implementations of PPCs on their respective Bayesian based method. The premise behind PPCs is evaluating the generated model on how well it is able to generate data similar to the data observed [3] utilizing in this case the B-splines dependent Dirichlet process and Bayesian normal linear regression model for the *AROC.bnp* and *AROC.bsp* objects respectively [14]. To exemplify these methods

we use the previously generated objects.

```
predictive.checks.AROC.bnp(bnpCribSex,
                           devnew = FALSE)
predictive.checks.AROC.bsp(bspCribSex,
                           devnew = FALSE)
```

While the mandatory argument is solely the AROC object, the *devnew* was changed from the default *TRUE* argument to display all depicted graphics on the same device. These graphics are histograms of the desired test statistics, by default minimum, maximum, median and skewness and Kernel density estimates showing diagnostic test outcome in the nondiseased group as well.

**Pooled ROC.** The package also provides two methods of polled ROC estimation, *pooledROC.emp* an implementation of empirical estimation proposed by Hsieh and Turnbull [5,14] and *pooledROC.BB* for Bayesian bootstrap estimation proposed by Gu et al. [4,14]. Both methods are similar to construct and have a similar output structure, mandatory arguments identify diseased and non-diseased groups in the test or marker column which can be achieved with a straightforward indexing in R.

```
# Pooled ROC methods

#------ BB method ------#
BBCrib <-
  pooledROC.BB(
    y0 = neonatal$crib[neonatal$status==0],
    y1 = neonatal$crib[neonatal$status==1])

plot(BBCrib)

#------ emp method ------#

EMPCrib <-
  pooledROC.emp(
    y0 = neonatal$crib[neonatal$status==0],
    y1 = neonatal$crib[neonatal$status==1])

plot(EMPCrib)
```

## 3   Results

All methods allowed for a visualization of their output with the standard *plot()* function, albeit a simple one, that allows for further customization if necessary.

The *summary()* output for AROC methods all follow a similar template, reporting the requested command with its respective arguments and the AAUC

of the resulting AROC curve. While this command offers no additional information for the AROC objects when compared to the *plot()*, that also displays the AAUC, the semiparametric frequentist method generates a specific element, *fit.h* that can be accessed and provides a direct statistical summary on the effect of the covariate to the curve and the fitted regression model (Fig. 1).

```
> summary(spCribSex$fit.h)

Call:
lm(formula = formula.healthy, data = data.h)

Residuals:
    Min      1Q  Median      3Q     Max
-2.0019 -1.9603 -0.9603  0.9981 12.9981

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.96026    0.11918  16.447   <2e-16 ***
sexMale      0.04164    0.16267   0.256    0.798
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.537 on 976 degrees of freedom
Multiple R-squared:  6.713e-05, Adjusted R-squared:  -0.00
09574
F-statistic: 0.06553 on 1 and 976 DF,  p-value: 0.798
```

**Fig. 1.** The *fit.h summary()* output for semiparametric frequentist adjustment of sex to the CRIB score.

Methods used to adjust the sex covariate to the CRIB ROC curve all displayed AAUC intervals - available through both the *plot()* and *summary()* functions, and curves that overlap each other, inferring no statistical differences between them, an expected yet important attribute, making all methods viable for use in this instance despite a seemingly slight overestimation by the bayesian nonparametric method that holds no statistical relevance. These three curves can be seen in Fig. 2.

The kernel and semiparametric frequentist method shared the same similarities between them when adjusting for the mother's age in the CRIB ROC score, showing both overlapping AAUCs and curves that infer no statistical difference despite the low bootstrap resamples used in the kernel method in the interest of time. Both curves can be seen in Fig. 3. Having established that all AROC adjustment methods display statistically similar curves what follows is to establish the same result for the unadjusted curve methods and then to compare the two. Both bayesian and empirical pooled ROC functions, seen in Fig. 4 displayed similar results just as the adjusted methods, a direct comparison between the two is displayed in Fig. 5C. A comparison between adjusted and unadjusted ROC curves for both the sex of the infant and mother's age can be seen in Fig. 5 A and B respectively. While the AAUC statistics in Fig. 2 through Fig. 4 show a clear overlap between their respective adjustments and the pooled ROC curve

and Fig. 5 show several intersections between the curves themselves, an additional source of information on the covariates' influence on their tests is present in the extended summary statistics of the $AROC.sp$ where p-value $= 0.798$ and p-value $= 0.9721$ is displayed for biological sex and mother's age respectively a figure well above the 0.05 required to entail any influence over the test for these covariates.
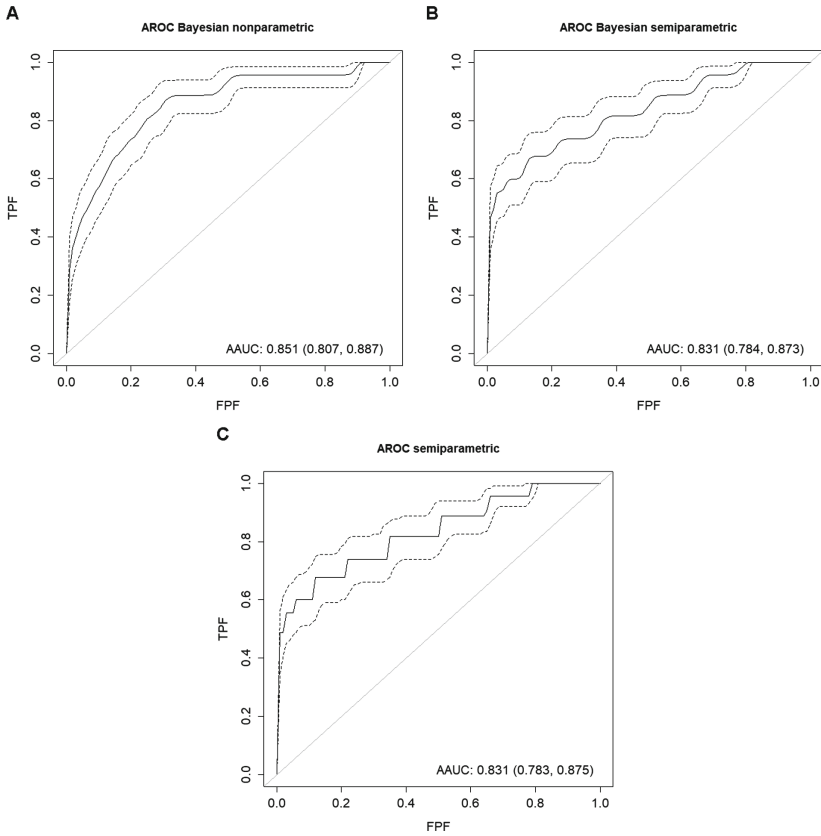


**Fig. 2.** AROC curve of CRIB score adjusted for the infants sex. **A**) AROC estimation with the nonparametric bayesian method. **B**) AROC estimation with the semiparametric bayesian method. **C**) AROC estimation with the semiparametric bayesian frequentist method.
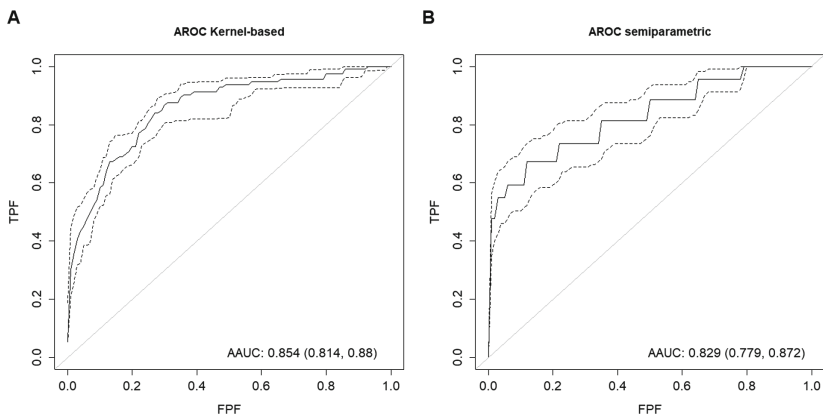
**Fig. 3.** AROC curve of CRIB score adjusted for the mother's age. **A**) AROC estimation with the kernel based method. **B**) AROC estimation with the semiparametric bayesian frequentist.
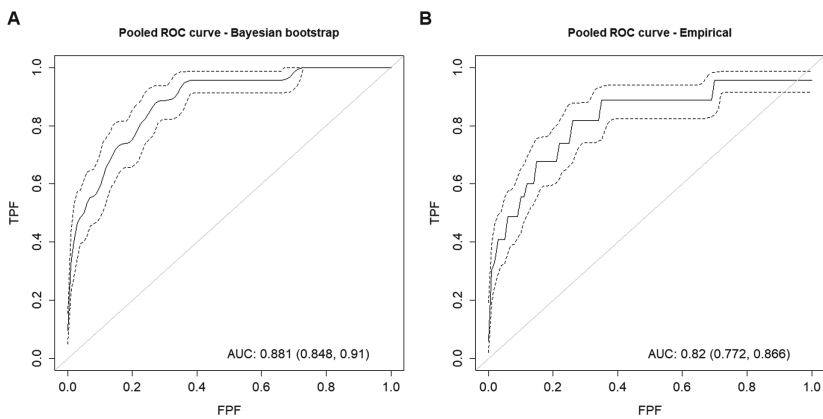


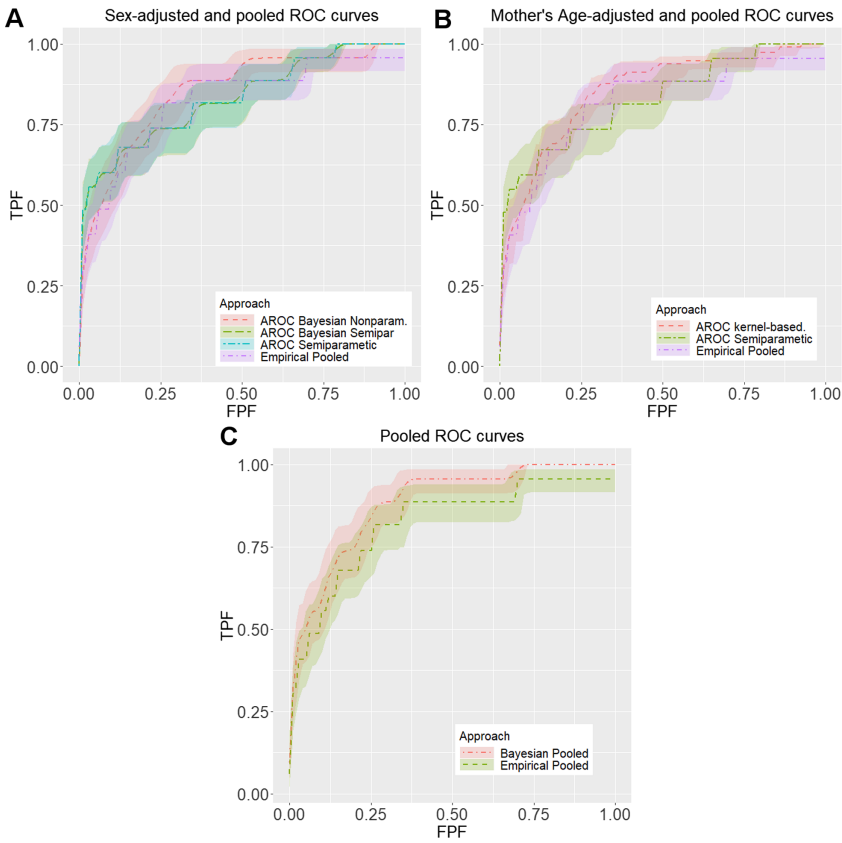**Fig. 4.** Pooled ROC curves for CRIB score. **A**) Bayesian bootstrap. **B**) Empirical estimation.

**Fig. 5.** Visual comparison between ROC and AROC curves. **A**) Comparison between AROC adjustment methods for biological sex and empirical pooled ROC curve. **B**) Comparison between AROC adjustment methods for mother's age and empirical pooled ROC curve. **C**) Visual Comparison between pooled ROC estimation methods

## 4   Discussion

All AROC estimation methods with the exception of AROC.kernel were capable of handling a categorical covariate rather than the standard continuous covariate. The visualization of the curves Fig. 2 through 5 provide insight on the behavior of the test and its associated covariate.

In efficiency, AROC.sp was the faster and most informative of the methods providing not only a visual representation of the AROC curve, but also descriptive test statistics inferring no statistical relevance to biological sex in the CRIB score that will be explored in the future.

Posterior predictive checks (PPCs) for non parametric bayesian method and for semiparametric bayesian method offer insight on their respective methods behavior and reliability that can lead to the choice of the most reliable method

between the two despite, as mentioned, in this case, no real difference being found between methods.

Pooled ROC methods deliver curves with small differences on AUC values due to the nature of their estimators however from AUC alone we can observe the AUC confidence intervals intercepting showing both methods deliver a statistical equivalent curve allowing analysts to choose a preferred method. We can expand this reasoning and see same applies to the all AROC curves further substantiating the AROC.sp statistics.

CRIB scores and biological sex show no clear relation for all methods tested, these findings while preliminary and supported by previous research results, [10,17] are in stark contrast with criticism to the scoring system that lead to the creation of the CRIB II [12]. This data will be further explored to analyze the confounding effect, if any, of this covariate.

## 5  Conclusion

The AROC package grants multiple methods of analysis and estimation of the AROC curve complemented with predictive checks and pooled ROC estimation. With the exception of the kernel based method, *AROC.kernel*, all AROC estimation methods have proven able to handle both categorical and continuous covariates and are even able to handle several sets of covariates of different natures. All methods generate AROC curves with no statistical differences, passing the choice of which method would be more suitable for the analysis directly to the analyst.

The reliability of this package will allow further research for possible covariates confounding CRIB score.

Overall the package finally brings to the R library a way to conduct a more thorough ROC and AROC analysis making it available for any R user.

## References

1. Brito, A.S.J.D., Matsuo, T., Gonzalez, M.R.C., de Carvalho, A.B.R., Ferrari, L.S.L.: CRIB score, birth weight and gestational age in neonatal mortality risk evaluation. Revista de saúde pública **37**(5), 597–602 (2003). http://www.ncbi.nlm.nih.gov/pubmed/14569335
2. Ezz- Eldin, Z.M., Abdel Hamid, T.A., Labib Youssef, M.R., Nabil, H.E.D.: Clinical risk index for babies (CRIB II) scoring system in prediction of mortality in premature babies. J. Clin. Diagnostic Res. **9**(6), SC08–SC11 (2015). https://doi.org/10.7860/JCDR/2015/12248.6012

3. Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., Gelman, A.: Visualization in Bayesian workflow. J. Royal Stat. Soc. Ser. A: Stat. Soc. **182**(2), 389–402 (2019). https://doi.org/10.1111/rssa.12378

4. Gu, J., Ghosal, S., Roy, A.: Bayesian bootstrap estimation of ROC curve. Stat. Med. **27**(26), 5407–5420 (2008). https://doi.org/10.1002/sim.3366, http://doi.wiley.com/10.1002/sim.3366

5. Hsieh, F., Turnbull, B.W.: Nonparametric and semiparametric estimation of the receiver operating characteristic curve. Ann. Statist. **24**(1), 25–40 (1996). https://doi.org/10.1214/aos/1033066197

6. Janes, H., Pepe, M.S.: Adjusting for covariates in studies of diagnostic, screening, or prognostic markers: an old concept in a new setting. Am. J. Epidemiol. **168**(1), 89–97 (2008). https://doi.org/10.1093/aje/kwn099

7. Janes, H., Pepe, M.S.: Adjusting for covariate effects on classification accuracy using the covariate-adjusted receiver operating characteristic curve. Biometrika **96**(2), 371–382 (2009). https://doi.org/10.1093/biomet/asp002

8. Lasko, T.A., Bhagwat, J.G., Zou, K.H., Ohno-Machado, L.: The use of receiver operating characteristic curves in biomedical informatics **38**, 404–415 (2005). https://doi.org/10.1016/j.jbi.2005.02.008

9. Machado, E., Costa, F., Braga, A.C.: Neonatalportugal. Mendeley Data (2019). https://doi.org/10.17632/jsmgcmfmdx.1

10. Mourão, M.F., Braga, A.C., Oliveira, P.N.: CRIB conditional on gender: nonparametric ROC curve. Int. J. Health Care Quality Assurance **27**(8), 656–663 (2014). https://doi.org/10.1108/IJHCQA-04-2013-0047

11. Park, S.H., Goo, J.M., Jo, C.H.: Receiver operating characteristic (ROC) curve: practical review for radiologists. Korean J. Radiol. **5**(1), 11 (2004). https://doi.org/10.3348/kjr.2004.5.1.11

12. Parry, G., Tucker, J., Tarnow-Mordi, W.O.: UK neonatal staffing study collaborative: CRIB II : an update of the clinical risk index for babies score For personal use. Only reproduce with permission from The Lancet Publishing Group. Lancet, **361** 1789–1791 (2003). https://doi.org/10.1016/S0140-6736(03)13397-1

13. Pepe, M.S.: The Statistical Evaluation of Medical Tests for Classification and Prediction (2003)

14. Rodriguez-Alvarez, M.X., Inacio de Carvalho, V.: AROC: Covariate-Adjusted Receiver Operating Characteristic Curve Inference (2018). https://CRAN.R-project.org/package=AROC, r package version 1.0

15. Rodríguez-Álvarez, M.X., Roca-Pardiñas, J., Cadarso-Suárez, C.: ROC curve and covariates: extending induced methodology to the non-parametric framework. Stat. Comput. **21**(4), 483–499 (2011). https://doi.org/10.1007/s11222-010-9184-1

16. Rodríguez-Álvarez, M.X., Roca-Pardiñas, J., Cadarso-Suárez, C., Tahoces, P.G.: Bootstrap-based procedures for inference in nonparametric receiver-operating characteristic curve regression analysis. Stat. Methods Med. Res. **27**(3), 740–764 (2018). https://doi.org/10.1177/0962280217742542

17. Terzic, S., Heljić, S.: Assessing mortality risk in very low birth weight infants. Medicinski arhiv **66**, 76–9 (2012). https://doi.org/10.5455/medarh.2012.66.76-79