

---

## Title

### ***Comparing the effectiveness of k-different treatments through the area under the ROC curve***

## Abstract

The area under the receiver-operating characteristic curve (AUC) has become a popular index not only for measuring the overall prediction capacity of a marker but also the strength of the association between continuous and binary variables. In the current considered study, the AUC was used for comparing the association size of four different interventions involving impulsive decision making, studied through an animal model, in which each animal provides several negative (pretreatment) and positive (posttreatment) measures. The problem of the full comparison of the average AUCs arises therefore in a natural way. We construct an analysis of variance (ANOVA) type test for testing the equality of the impact of these treatments measured through the respective AUCs and considering the random-effect represented by the animal. The use (and development) of a post hoc Tukey's HSD-type test is also considered. We explore the finite-sample behaviour of our proposal via Monte Carlo simulations, and analyse the data generated from the original problem. An R package implementing the procedures is provided in the supporting information. © 2024 John Wiley & Sons Ltd.

## Authors

Martínez-Cambor P.; Pérez-Fernández S.; Dwiel L.L.; Doucette W.T.

## Author full names

---

Martínez-Cambor, Pablo (24462229000); Pérez-Fernández, Sonia (56490868100);  
Dwiel, Lucas L. (56096054800); Doucette, Wilder T. (25822476100)

**Author(s) ID**

24462229000; 56490868100; 56096054800; 25822476100

**Year**

2024

**Source title**

Stat

**Volume**

13.0

**Issue**

2

**Art. No.**

e672

---

## DOI

10.1002/sta4.672

## Link

<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85189641081&doi=10.1002%2fst4.672&partnerID=40&md5=7f76ca32f76e35388c306413e2637388>

## Affiliations

Department of Anesthesiology, Geisel School of Medicine at Dartmouth, Hanover, NH, United States; Faculty of Health Sciences, Universidad Autonoma de Chile, Providencia, Chile; Department of Statistics and Operations Research, University of Oviedo, Asturias, Spain; Department of Psychiatry, Geisel School of Medicine at Dartmouth, Hanover, NH, United States

## Authors with affiliations

Martínez-Cambor P., Department of Anesthesiology, Geisel School of Medicine at Dartmouth, Hanover, NH, United States, Faculty of Health Sciences, Universidad Autonoma de Chile, Providencia, Chile; Pérez-Fernández S., Department of Statistics and Operations Research, University of Oviedo, Asturias, Spain; Dwiell L.L., Department of Psychiatry, Geisel School of Medicine at Dartmouth, Hanover, NH, United States; Doucette W.T., Department of Psychiatry, Geisel School of Medicine at Dartmouth, Hanover, NH, United States

---

## Author Keywords

ANOVA test; area under the curve; brain quantification; post hoc test; random-effects

## Funding Details

National Institute of Mental Health, NIMH; Asturias Government, (PID2020-118101GB-I00); Ministerio de Ciencia e Innovación, MCIN, (1R01MH134141)

## Funding Texts

Funding text 1: This work was supported by Grant GRUPIN AYUD/2021/50897 from the Asturias Government, by Grant PID2020-118101GB-I00 from Ministerio de Ciencia e Innovación (Spanish Government), and by Grant 1R01MH134141 from the National Institute of Mental Health.; Funding text 2: This work was supported by Grant GRUPIN AYUD/2021/50897 from the Asturias Government, by Grant PID2020-118101GB-I00 from Ministerio de Ciencia e Innovación (Spanish Government), and by Grant 1R01MH134141 from the National Institute of Mental Health.

## References

Bandos A.I., Rockette H.E., Gur D., A permutation test sensitive to differences in areas for comparing ROC curves from a paired design, *Statistics in Medicine*, 24, 18, pp. 2873-2893, (2005); Buzsaki G., Anastassiou C., Koch C., The origin of extracellular fields and currents: EEG, ECoG, LFP and spikes, *Nature Review*

---

Neurosciences, 13, 6, pp. 407-420, (2012); de Hond A.A.H., Steyerberg E.W., van Calster B., Interpreting area under the receiver operating characteristic curve, Lancet Digital Health, 4, 12, pp. e853-e855, (2022); DeLong E.R., DeLong D.M., Clarke-Pearson D.L., Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach, Biometrics, 44, 3, pp. 837-845, (1988); Demidenko E., The p-value you can't buy, American Statistician, 1, 70, pp. 33-38, (2016); Dixon W.J., Simplified estimation from censored normal samples, The Annals of Mathematical Statistics, 31, 2, pp. 385-391, (1960); Dorfman D.D., Berbaum K.S., Metz C., Receiver operating characteristic rating analysis: Generalization to the population of readers and patients with the jackknife method, Investigative Radiology, 27, 9, pp. 723-731, (1992); Dwiel L.L., Khokhar J.Y., Connerney M.A., Green A.I., Doucette W.T., Finding the balance between model complexity and performance: Using ventral striatal oscillations to classify feeding behavior in rats, PLOS Computational Biology, 15, 4, pp. 1-19, (2019); Friedman J., Tibshirani R., Hastie R., Regularization paths for generalized linear models via coordinate descent, Journal of Statistical Software, 33, 1, pp. 1-22, (2010); Gallas B.D., Bandos A., Samuelson F.W., Wagner R.F., A framework for random-effects ROC analysis: Biases with the bootstrap and other variance estimators, Communications in Statistics - Theory and Methods, 38, 15, pp. 2586-2603, (2009); Goeman J.J., Solari A., Comparing three groups, The American Statistician, 76, 2, pp. 168-176, (2022); Hampel F.R., The influence curve and its role in robust estimation, Journal of the American Statistical Association, 69, 346, pp. 383-393, (1974); Hanley J.A., McNeil B.J., The meaning and use of the area under a receiver operating characteristic (ROC) curve, Radiology, 143, 1, pp. 29-36, (1982); Harris B.A.L., Uhelski M.L., Li A.L., Extrapolating meaning from local field potential recordings, Journal of Integrative Neuroscience, 16, 1, pp. 107-126, (2017); Hillis S.L., Obuchowski N.A., Scharz K.M., Berbaum K.S., A comparison of the Dorfman-Berbaum-Metz and Obuchowski-Rockette methods for receiver operating

---

characteristic (ROC) data, *Statistics in Medicine*, 24, 10, pp. 1579-1607, (2005); Holm K., Christman N.J., Post hoc tests following analysis of variance, *Research in Nursing & Health*, 8, 2, pp. 207-210, (1985); Hsieh F., Turnbull B.W., Nonparametric and semiparametric estimation of the receiver operating characteristic curve, *The Annals of Statistics*, 24, 1, pp. 25-40, (1996); Janssens A., Cecile J.W., Martens F.K., Reflection on modern methods: Revisiting the area under the ROC Curve, *International Journal of Epidemiology*, 49, 4, pp. 1397-1403, (2020); Kacker R.N., Combining information from interlaboratory evaluations using a random effects model, *Metrologia*, 41, 3, pp. 132-136, (2004); Kim E., Zhang Z., Wang Y., Zeng D., Power calculation for comparing diagnostic accuracies in a multi-reader, multi-test design, *Biometrics*, 70, 4, pp. 1033-1041, (2014); Martinez-Cambor P., Perez-Fernandez S., Diaz-Coto S., The area under the ROC curve as effect size measure, *Pakistan Journal of Statistics*, 36, 1, pp. 13-28, (2020); Mitchell M.R., Potenza M.N., Recent insights into the neurobiology of impulsivity, *Current Addiction Reports*, 1, 4, pp. 309-319, (2014); Moeller F.G., Barratt E.S., Dougherty D.M., Schmitz J.M., Swann A.C., Psychiatric aspects of impulsivity, *American Journal of Psychiatry*, 158, 11, pp. 1783-1793, (2001); Nakas C., Bantis L., Gatsonis C., ROC analysis for classification and prediction in practice, (2023); Obuchowski N.A., Nonparametric analysis of clustered ROC curve data, *Biometrics*, 53, 2, pp. 567-578, (1997); Obuchowski N.A., Rockette H.E., Hypothesis testing of diagnostic accuracy for multiple readers and multiple tests an ANOVA approach with dependent observations, *Communications in Statistics - Simulation and Computation*, 24, 2, pp. 285-308, (1995); Pepe M.S., The statistical evaluation of medical tests for classification and prediction, (2003); Shu D., Zou G., Revisiting sample size planning for receiver operating characteristic studies: A confidence interval approach with precision and assurance, *Statistical Methods in Medical Research*, 32, 4, pp. 748-759, (2023); Tcheuko L., Gallas B., Samuelson F., Using ANOVA/random-effects variance estimates to compute a two-sample U-statistic of order (1,1) estimate of

---

---

variance, *Journal Statistical Theory Practice*, 10, pp. 87-99, (2016); Tibshirani R., Regression shrinkage and selection via the lasso, *Journal of the Royal Statistical Society. Series B (Methodological)*, 58, 1, pp. 267-288, (1996); Tukey J., Comparing individual means in the analysis of variance, *Biometrics*, 5, 2, pp. 99-114, (1949); van der Vaart A.W., *Asymptotic statistics*, (2000); Zhou X.-H., Obuchowski N.A., McClish D.K., *Statistical methods in diagnostic medicine*, (2002)

## **Correspondence Address**

P. Martínez-Cambor; Department of Anesthesiology, Geisel School of Medicine at Dartmouth, Hanover, 7 Lebanon Street, Suite 309, Hinman Box 7261, 03751, United States; email: Pablo.Martinez-Cambor@hitchcock.org

## **Publisher**

John Wiley and Sons Inc

## **ISSN**

20491573

## **Language of Original Document**

English

## **Abbreviated Source Title**

Stat

---

## Document Type

Article

## Publication Stage

Final

## Source

Scopus

## EID

2-s2.0-85189641081