

iMRMC: Software to do Multi-reader Multi-case Statistical Analysis of Reader Studies

Catalog of Regulatory Science Tools to Help Assess New Medical Devices

Technical Description

The primary objective of the iMRMC statistical software is to assist investigators with analyzing and sizing multi-reader multi-case (MRMC) reader studies that compare the difference in the area under Receiver Operating Characteristic curves (AUCs) from two modalities. The [iMRMC application](#) is a software package that includes simulation tools to characterize bias and variance of the MRMC variance estimates.

The core elements of this application include the ability to perform MRMC variance analysis and the ability to size an MRMC trial.

- The core [iMRMC application](#) is a stand-alone, precompiled, license-free Java applications and the source code. It can be used in GUI mode or on the command line.
- There is also an [R package](#) that utilizes the core Java application. Examples for using the programs can be found in the R help files.
- Additional functionality of the [GitHub package](#) includes an example to guide users on how to perform a noninferiority study using the iMRMC R package.

The software treats arbitrary study designs that are not "fully-crossed."

Intended Purpose

The iMRMC package analyzes data from Multiple Readers and Multiple Cases (MRMC) studies, which are often imaging studies where clinicians (readers) evaluate patient images (cases). The MRMC methods apply to any scenario in which clinicians interpret data to make clinical decisions. The iMRMC package calculates the reader-averaged area under the receiver operating characteristic curve: the AUC of the ROC curve. AUC is a diagnostic performance measure. Additional functions analyze other endpoints (binary performance and score differences). This package also estimates variances, confidence intervals and p-values. These uncertainty characteristics are needed for hypothesis tests to size and assess the efficacy of diagnostic imaging devices and computer aids (artificial intelligence)

The analysis is important because imaging studies are designed so that every reader reads every case in all modalities, a fully-crossed study. In this case, the data is cross-correlated, and the readers and cases are considered to be cross-correlated random effects. An MRMC analysis accounts for the variability and correlations from the readers and cases when estimating variances, confidence intervals, and p-values. The functions in this package can treat arbitrary study designs and studies with missing data, not just fully-crossed study designs.

The methods in the iMRMC package are not standard. The package permits industry statisticians to use a validated statistical analysis method without having to develop and validate it themselves.

Related FDA Product Codes

The FDA product codes this tool is applicable to include, but are not limited to:

- KPS: System, Tomography, Computed, Emission
- LLZ: System, Image Processing, Radiological
- PAA: Automated Breast Ultrasound
- POK: Computer-Assisted Diagnostic Software For Lesions Suspicious For Cancer
- QDQ: Radiological Computer Assisted Detection/Diagnosis Software For Lesions Suspicious For Cancer
- QPN: Software Algorithm Device To Assist Users In Digital Pathology
- QNP: Gastrointestinal lesion software detection system

Testing

The tool has been characterized through simulations (bias and variance of the estimates) and has been compared with other methods as appropriate for the task.

The following peer-reviewed research includes the detailed verification methods and results

- Gallas, B. D., Chen, W., Cole, E., Ochs, R., Petrick, N., Pisano, E. D., Sahiner, B., Samuelson, F. W., & Myers, K. J. (2019). Impact of prevalence and case distribution in lab-based diagnostic imaging studies. *Journal of Medical Imaging*, 6(1), 015501. <https://doi.org/10.1117/1.JMI.6.1.015501>
 - Desc: Study that uses the software and related research methods and study designs in a large study. Supplementary materials include data and scripts to reproduce study results.

- Gallas, B. D. (2006). One-shot estimate of MRMC variance: AUC. *Acad Radiol*, 13(3), 353–362. <https://doi.org/10.1016/j.acra.2005.11.030>
 - Desc: Original description of method and validation with simulations. Results comparable to jackknife resampling technique.
- Gallas, B. D., Pennello, G. A., & Myers, K. J. (2007). Multireader multisequence variance analysis for binary data. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, 24(12), B70-80. <https://doi.org/10.1364/josaa.24.000b70>
 - Generalize method to binary performance measures.
- Gallas, B. D., Bandos, A., Samuelson, F., & Wagner, R. F. (2009). A framework for random-effects ROC analysis: Biases with the bootstrap and other variance estimators. *Commun Stat A-Theory*, 38(15), 2586–2603. <https://doi.org/10.1080/03610920802610084>
 - Provide framework for understanding method and comparing to other methods analytically and with simulations.
- Gallas, B. D., & Brown, D. G. (2008). Reader studies for validation of CAD systems. *Neural Networks Special Conference Issue*, 21(2), 387–397. <https://doi.org/10.1016/j.neunet.2007.12.013>

Limitations

Currently, the tool can produce negative variance estimates if the relevant dataset is small.

Supporting Documentation

Tool websites:

- Primary: <https://github.com/DIDSR/iMRMC>
- Secondary: <https://cran.r-project.org/web/packages/iMRMC/index.html>

User manual for java app

- http://didsr.github.io/iMRMC/000_iMRMC/userManualPDF/iMRMCUserManual.pdf

User manual for R package

- <https://cran.r-project.org/web/packages/iMRMC/iMRMC.pdf>

FAQs

- <https://github.com/DIDSR/iMRMC/wiki/iMRMC-FAQ>

Supplementary materials

- [Data and scripts to reproduce results for manuscripts that use iMRMC](#)
- <https://github.com/DIDSR/iMRMC/wiki/iMRMC-Datasets>

Related Work

- Chen, W., Gong, Q., Gallas, B.D. (2018). Paired split-plot designs of multireader multicase studies. *Journal of Medical Imaging* 5, 031410. <https://doi.org/10.1117/1.JMI.5.3.031410>
- Obuchowski, N.A., Gallas, B.D., Hillis, S.L. (2012). Multi-Reader ROC studies with Split-Plot Designs: A Comparison of Statistical Methods. *Acad Radiol* 19, 1508– 1517. <https://doi.org/10.1016/j.acra.2012.09.012>
- Gallas, B.D., Chan, H.-P., D’Orsi, C.J., Dodd, L.E., Giger, M.L., Gur, D., Krupinski, E.A., Metz, C.E., Myers, K.J., Obuchowski, N.A., Sahiner, B., Toledano, A.Y., Zuley, M.L. (2012). Evaluating imaging and computer-aided detection and diagnosis devices at the FDA. *Acad Radiol* 19, 463–477. <https://doi.org/10.1016/j.acra.2011.12.016>
- Obuchowski, N. A., Gallas, B. D., & Hillis, S. L. (2012). Multi-Reader ROC studies with Split-Plot Designs: A Comparison of Statistical Methods. *Academic Radiology*, 19(12), 1508–1517. <https://doi.org/10.1016/j.acra.2012.09.012>
- Gallas, B. D., & Hillis, S. L. (2014). Generalized Roe and Metz ROC model: Analytic link between simulated decision scores and empirical AUC variances and covariances. *J Med Img*, 1(3), 031006. <https://doi.org/doi:10.1117/1.JMI.1.3.031006>

Contact

- RST_CDRH@fda.hhs.gov

Tool Reference

In addition to citing relevant publications please reference the use of this tool using DOI: 10.5281/zenodo.8383591

For more information:

- [Catalog of Regulatory Science Tools to Help Assess New Medical Devices](#)