Recommended software/packages for meta-analysis of diagnostic accuracy

Junfeng Wang[1], Mariska Leeflang[2]

[1]Department of Biomedical Data Sciences, Leiden University Medical Center, The Netherlands; [2]Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Academic Medical Centre, The Netherlands

Correspondence to: Junfeng Wang, PhD. Statistician, Department of Biomedical Data Sciences, Leiden University Medical Center, Einthovenweg 20, 2333ZC Leiden, The Netherlands. Email: j.wang@lumc.nl.

Abstract: When performing a meta-analysis of diagnostic accuracy, authors usually need to use a statistical package. Many different packages are available but not all of them are well developed. In this review, we first make an inventory of most popular packages providing meta-analysis of diagnostic accuracy, and then evaluate them by their accessibility, usability, and functions and methodology, and give an overall assessment for each package. This methodological review will enable review authors to understand what the packages can do for them and to choose the correct methods.

Keywords: Diagnostic test; meta-analysis; HSROC; bivariate model; systematic review

Introduction

It has been almost 30 years since the publication of the first meta-analysis of diagnostic test accuracy (DTA) (1). Since then, the statistical methods evolved from simply following the approaches used for intervention meta-analyses to the summary ROC (sROC) model (also known as Moses-Littenberg model) which takes in to account the threshold effect (2), and then to more advanced models, i.e. bivariate model (3) and HSROC model (4), which also allow for both within and between study variability.

Although HSROC and bivariate models are commonly regarded as the best practice and recommended in the Cochrane Handbook (5), suboptimal methods are still frequently used in studies performing meta-analysis of diagnostic accuracy, since they are widely implemented in software and packages and more accessible than the preferred ones. In the recent book of Diagnostic Meta-Analysis, Doebler *et al.* gave an overview of packages available for DTA meta-analysis and briefly discussed their strength and weakness (6). Any statistical software with general mixed model can be employed for fitting the advanced meta-analysis models. However, researchers prefer packages designed for specific analysis and without extra coding work. So in this review, we will make an inventory of most popular packages providing meta-analysis of diagnostic accuracy, and evaluate them by their accessibility, usability, and functions and methodology with a traffic light approach (*Table 1*).

By bridging the methodology gap between statisticians and clinicians, we enable review authors to understand what the packages can do for them and to choose the correct methods.

Meta-DiSc

Accessibility 🔴 ⚪ ⚪

Meta-DiSc (7) is a freeware software developed by a group of statisticians in Spain to perform meta-analysis of diagnostic and screening tests. It used to be available for download from the group's website (http://www.hrc.es/investigacion/metadisc_en.htm), and now the new version (Meta-DiSc 2.0) is under development. The old version (Meta-DiSc 1.4) is removed from the website, but can still be found if searching on the internet.

Table 1 Criterion and implication of the traffic lights

Criterion	🔴	🟡	🟢
Accessibility	Not accessible	Commercial software	Free software
Usability	Very difficult to use	Difficult to use	Easy to use
Functions for meta-analysis and methodology soundness	Methods are not preferred	Functions are not complete	Preferred methods are available
Overall	Not recommended	Neutral	Recommended

Usability ⚪ ⚪ 🟢

Meta-DiSc has a user-friendly interface with drop-down menus, dialog boxes and online help files which include a user manual and a description of the implemented statistical methods.

Data from primary studies, including study ID, accuracy data (TP, FP, TN, FN) and study level co-variates, can be entered into the datasheet by either typing, copy-pasting or importing.

Functions for meta-analysis and methodology soundness 🔴 ⚪ ⚪

Meta-DiSc can produce forest plots of sensitivities, specificities, likelihood ratios (LRs) and diagnostic odds ratios (DOR) and also ROC plane scatter-plots. It performs statistical pooling of sensitivity, specificity, LRs, DOR using fixed effects model (Mantel-Haenszel method) and random effects models (DerSimonian Laird method). Meta-DiSc can fit sROC curve with Littenberg and Moses model, and estimate AUC and the Q* index. Besides these basic analyses, Meta-DiSc also allows exploration of heterogeneity (Chi-square, Cochran-Q and I-squared), and meta-regression.

However, separate pooling of sensitivities and specificities and the Littenberg and Moses model have some inherent statistical shortcomings, as the between-study variance is not included. More advanced methods (i.e., HSROC and bivariate model) are not implemented. Last but not least, Meta-DiSc provides a bunch of analyses, but not all of them are suitable for meta-analysis of test accuracy studies.

Overall 🔴 ⚪ ⚪

Meta-DiSc is free for all users and easy to learn and use. However, just as the authors stated, the current version of Meta-DiSc uses outdated statistical methods and should

not be used for making inferences. Their contribution and efforts preventing reviewers from using the not recommended methods are highly appreciated, so let's look forward to Meta-DiSc 2.0.

RevMan 5

Accessibility ⚪ ⚪ 🟢

Review Manager 5 (RevMan 5) (8) is Cochrane's software for preparing and maintaining systematic reviews, as well as performing meta-analysis. It can be downloaded freely from the Cochrane website (https://community.cochrane.org/help/tools-and-software/revman-5). The development team are currently working on the next generation of online review-writing software, RevMan Web, and the current version (RevMan 5.3) is no longer being developed, but Cochrane authors can still get support when using the programme.

Usability ⚪ 🟡 ⚪

RevMan is designed to facilitate all procedures in systematic reviews, including text, characteristics of studies, comparison tables, study data input and management, meta-analysis of the data entered, and graphical presentation of the results. A very detailed User Guide and many tutorials are provided by Cochrane. However, most users think RevMan is not very user friendly.

Functions for meta-analysis and methodology soundness ⚪ 🟡 ⚪

RevMan only provides limited analyses for diagnostic accuracy studies. Sensitivities and specificities are separately plotted in forest plot only for exploring and presenting purpose, no summary points or heterogeneity measures are provided (in some extent, it is good). Heterogeneity

is assessed by stratified analysis of studies and separate ROC curves per subgroup, but no formal testing (e.g., P value) will be calculated to determine whether accuracy is significantly different between subgroups.

More complex models, such as the HSROC and Bivariate model, cannot be fitted in RevMan. These analyses can be performed in external statistical packages (e.g., SAS, Stata) and then import the parameters needed to draw the summary ROC curves into RevMan. RevmMan does not present the actual numbers of the output, only uses it for illustration. Furthermore, the parameters requested come from STATA or SAS. If someone uses R or another software, it requires some more background knowledge to choose or translate the right parameters.

Overall ◯ 🟡 ◯

RevMan is endorsed by Cochrane and its quality is ensured by leading research groups in this area. It can help the users with performing the meta-analysis and writing a systematic review step by step. The implementation of advanced models in RevMan is possible but not straight forward.

Stata

Accessibility ◯ 🟡 ◯

Stata (9) is a statistical package widely used in the fields of biomedicine and epidemiology, as well as economics, sociology and political science. Stata is a commercial software, and license needs to be purchased.

Usability ◯ ◯ 🟢

Stata is considered as a "Fast, Accurate, Easy to use" integrated software package. It has both a command-line interface and menus and dialog boxes for all built-in commands. Stata allows user-written commands, which can be installed directly in Stata ("ssc install name_command") or downloaded from the internet (ado-files).

Functions for meta-analysis and methodology soundness ◯ ◯ 🟢

There are two user-written commands in Stata, i.e., midas (10) and metandi (11), which implement some contemporary statistical methods for meta-analysis of binary diagnostic test.

In midas, summary sensitivity and specificity are derived from the bivariate model, and can be depicted in SROC space with the confidence and prediction regions (optional). Midas can also facilitate explore of heterogeneity and publication bias. However, due to the threshold effect, I-squared (in the forest plot) and funnel plot of log(DOR) are not as meaningful as in intervention meta-analysis. The main advantages of midas are the incorporation of meta-regression and graphical model checking, which are usually missing in other packages.

Metandi uses the same bivariate method as midas, but has less analyses on heterogeneity (again, which is not bad). The greatest contribution of metandi is that it provides the estimates of parameters in HSROC model (Λ, Θ, β, σ_α^2, σ_θ^2) and bivariate model (μ_A, μ_B, σ_A^2, σ_B^2, ρ_{AB}). When review authors perform a Cochrane systematic review, these parameters feed the drawing of the Figures in RevMan 5.

Overall ◯ ◯ 🟢

Stata can provide parameter estimates of both HSROC and bivariate models, as well as other analyses. Given Stata is familiar to clinical researchers and the click-to-run feature, we give our recommendation to it.

SAS

Accessibility ◯ 🟡 ◯

SAS is short for its previous name "Statistical Analysis System". SAS gains its popularity in pharmaceutical and financial industries with its great strength in managing complex data sets. SAS is also extremely expensive, but the good news is SAS launched a University Edition which is free for academic and non-commercial use.

Usability ◯ 🟡 ◯

SAS provides extraordinary range of data analysis tools for users, however it is difficult to use and learn and considered as power users' choice. There are few statistical analysis you cannot do in SAS, but you need to first manage using SAS and the learning curve is steep. User designed algorithm can be coded as SAS Macro, which is similar to ado-file in Stata or R package. SAS Macro aims to make complex algorithm user-friendly, but in the end you will find it is not user-friendly at all.

For DTA meta-analysis, a SAS macro named MetaDAS

was developed to automate the fitting of bivariate and HSROC models (12). A detailed user guild accomplished by a quick reference and worked example is available for users not very familiar with SAS.

Functions for meta-analysis and methodology soundness ◯ ◯ ●

MetaDAS can fit hierarchical models for analysis of a single test, test comparisons (by including test type as a covariate), investigations of heterogeneity and meta regression with study level covariates. Parameters of the HSROC model (Λ, Θ, β, σ_α^2, σ_θ^2) and bivariate model (μ_A, μ_B, σ_A^2, σ_B^2, ρ_{AB}) are estimated and summary estimates of test accuracy measures (sensitivity, specificity, DOR, LR+ and LR-) of single tests and relative measures of test accuracy in test comparison are provided as well. Covariates can be added the both models, and more than one covariate is allowed, which is the main advantage of SAS macro MetaDAS compared to other packages like midas in Stata and mada in R.

Overall ◯ ● ◯

SAS is a powerful tool for almost all statistical analyses, and the MetaDAS macro is written in high quality. However, SAS is expensive and not every institute has a SAS license. This also leads to many clinical researchers don't have experience with using SAS.

R

Accessibility ◯ ◯ ●

R is a free software environment for statistical analysis and most popular among statisticians. R can be downloaded from CRAN (Comprehensive R Archive Network) with mirrors all over the world. There are several R packages can facilitate DTA meta-analysis, which also can be found in CRAN or installed directly from R by using install. packages() commend.

Usability ◯ ● ◯

Fairly speaking, R is not very user friendly. The command line can be quite daunting to a beginner of R. Some graphical user interfaces for R, e.g. RStudio, may help a bit, but they are still not as convenient as those click-to-run software. Most R packages provide detailed explanatory

document with examples for all the functions in the package.

Functions for meta-analysis and methodology soundness ◯ ◯ ●

One advantage of R is that new methods can be implemented very quickly in user written R packages. The disadvantage may be that these methods and packages were not widely discussed or validated before being online.

Many packages were developed for DTA meta-analysis studies, including mada (13), packages using Baysesian approach [HSROC (14), bamdit (15), meta4diag (16)], and packages for special situations [DiagMeta (17) for several thresholds from one study, Metatron (18) for imperfect reference standard]. In this assessment, we focus on mada, which provides the standard approach and is more comparable to other software.

Mada fits the bivariate model and then transforms the bivariate model parameters to HSROC estimates. It also calculates the area under the summary ROC curve as a single accuracy measure, however it is shown in recent research this measure may lead to biased results (19). Meta-regression can be performed one variable at each time. Mada also provides other not widely used approaches, e.g., univariate pooling of DOR, and Proportional hazards model approach.

Overall ◯ ◯ ●

The wide variety of packages in R give the researcher more options when performing the analysis, while also bringing some difficulties in decision making which package to use. We recommend mada to basic users, and other packages may be interesting to statisticians who are doing methodological research.

Summary

Most clinician review authors rely on the software when they implement statistical methods, thus the methodological quality of the review heavily depends on the package they choose. In this evaluation, we give an overview of packages available for meta-analysis of diagnostic accuracy.

Table 2 shows the summary results of the evaluation. R and Stata are the two recommended packages for their easy to use and methodology. RevMan 5 and SAS are very useful when doing Cochrane DTA review. Researchers should stop

Table 2 Evaluation of packages for meta-analysis of diagnostic accuracy

Criterion	Meta-DiSc	RevMan 5	Stata	SAS	R
Accessibility	🔴	🟢	🟡	🟡	🟢
Usability	🟢	🟡	🟢	🟡	🟡
Functions for meta-analysis and methodology soundness	🔴	🟡	🟢	🟢	🟢
Overall	🔴	🟡	🟢	🟡	🟢

using Meta-DiSc in its current version.

Acknowledgments

Footnote

Provenance and Peer Review: This article was commissioned by the editorial office, *Journal of Laboratory and Precision Medicine* for the series "Statistics Corner". The article has undergone external peer review.

Conflicts of Interest: J Wang has no conflicts of interest to declare; M Leeflang is the co-convenor of the Cochrane Screening and Diagnostic Test Methods Group and Editor of the Cochrane Collaboration's Diagnostic Test Accuracy Editorial Team.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Leeflang MM, Deeks JJ, Takwoingi Y, et al. Cochrane diagnostic test accuracy reviews. Syst Rev 2013;2:82.
2. Moses LE, Shapiro D, Littenberg B. Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. Stat Med 1993;12:1293-316.
3. Reitsma JB, Glas AS, Rutjes AW, et al. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol 2005;58:982-90.
4. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. Stat Med 2001;20:2865-84.
5. Macaskill P, Gatsonis C, Deeks J, et al. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy. Version 09. London: The Cochrane Collaboration, 2010.
6. Doebler P, Bürkner P-C, Rücker G. Statistical Packages for Diagnostic Meta-Analysis and Their Application. Diagnostic Meta-Analysis. Springer, 2018:161-81.
7. Zamora J, Abraira V, Muriel A, et al. Meta-DiSc: a software for meta-analysis of test accuracy data. BMC Med Res Methodol 2006;6:31.
8. Collaboration NCCTC. Review manager (RevMan) [computer program] Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration 2014.
9. StataCorp L. Stata Statistical Software: Release 15 College Station, TX, 2017.
10. Dwamena B, Sylvester R, Carlos R. midas: Stata module for meta-analytical integration of diagnostic test accuracy studies. Statistical Software Components S456880, Department of Economics, Boston College. 2007.
11. Harbord RM, Whiting P. Metandi: meta-analysis of diagnostic accuracy using hierarchical logistic regression. The Stata Journal 2009;9:211-29.
12. Takwoingi Y, Deeks J. editors. METADAS: a SAS macro for meta-analysis of diagnostic accuracy studies. Methods for evaluating medical tests symposium. University of Birmingham; 2008: Citeseer.
13. Doebler P, Holling H. Meta-analysis of diagnostic accuracy with mada. Available online: https://cran.r-project.org/web/packages/mada/vignettes/mada.pdf
14. Schiller I, Dendukuri N. HSROC: An R package for Bayesian meta-analysis of diagnostic test accuracy. 2015.
15. Verde PE. bamdit: An R package for Bayesian meta-

analysis of diagnostic test data. Journal of Statistical Software, Articles 2018;86:1-32.

16. Guo J, Riebler A. meta4diag: Bayesian bivariate meta-analysis of diagnostic test studies for routine practice. arXiv preprint arXiv:151206220 2015.

17. Steinhauser S, Schumacher M, Rücker G. Modelling multiple thresholds in meta-analysis of diagnostic test accuracy studies. BMC Med Res Methodol 2016;16:97.

18. Huang H. Metatron: meta-analysis for classification data and correction to imperfect reference. R package version 0.1-1. 2014. 2018.

19. Wang J, Zwinderman AH, Leeflang MM. The area under the ROC curve may be a biased performance measure for meta-analysis of diagnostic accuracy studies. A simulation study. Global Evidence Summit; Cape Town, South Africa 2017.