Journal Title
XX(X):1–3

DOI: 10.1177/ToBeAssigned
www.sagepub.com/

Comment on 'A vine copula mixed effect model for trivariate meta-analysis of diagnostic test accuracy studies accounting for disease prevalence' by Aristidis K Nikoloulopoulos

Annika Hoyer[1] and Oliver Kuss[1]

With interest and excitement we noticed the recently published paper from Nikoloulopoulos[1]. The author proposes an approach for the meta-analysis of diagnostic accuracy studies accounting for disease prevalence with modelling random effects by vine copulas. The use of copulas in meta-analysis opens up a new research area and we expect interesting and practically useful results to emerge in the future. For example, Nikoloulopoulos extends the family of available copulas to asymmetric ones and examines all possible permutations of the vine copulas.

In his paper, Nikoloulopoulos repeatedly refers to recent works of ours[2], henceforth referred as the HK model, and we like to comment on some aspects.

[1] German Diabetes Center, Leibniz Institute for Diabetes Research at Heinrich Heine University Düsseldorf, Institute for Biometry and Epidemiology

Corresponding author:
Annika Hoyer, Deutsches Diabetes-Zentrum, Institut für Biometrie und Epidemiologie, Auf'm Hennekamp 65, 40225 Düsseldorf, Germany
Email: annika.hoyer@ddz.uni-duesseldorf.de

Prepared using sagej.cls *[Version: 2015/06/09 v1.01]*

Nikoloulopoulos[1] explicitly states twice in his paper that the main parameters of interest in the meta-analysis of diagnostic studies are sensitivity and specificity and we fully agree. However, we were surprised that Nikoloulopoulos calls our model inefficient, because in one of his previous papers[3] he observed that the HK model had problems with estimating the copula association parameter in the bivariate case. Again, the parameters of interest are sensitivity and specificity, and we show in our paper[2] that these can be estimated also in the trivariate case with essentially no bias and good coverage from the HK model. Moreover, results from the HK model are frequently better than those from the standard trivariate generalized linear mixed model (GLMM) as proposed by Chu et al.[4]

As an important advantage of the HK, as compared to the GLMM model, we consider its robustness. Our SAS NLMIXED code for the copula models showed superior convergence as compared to PQL estimation (SAS PROC GLIMMIX) in the bivariate as well as in the trivariate case. We even abandoned estimating parameters by Gauss-Hermite-Quadrature estimation for the GLMM model (SAS PROC NLMIXED) in the course of our trivariate simulation study because of numerical instabilities. This is certainly to be expected because fitting the HK model reduces to simply fitting a trivariate distribution, whereas estimating parameters from GLMMs requires evaluating or approximating complicated integrals in higher dimensions. Unfortunately, Nikoloulopoulos[1] does not report on convergence issues from his proposed Gauss-Legendre algorithm, which is also a numerical integration method prone to similar convergence problems. It would be thus very interesting to learn how the author's model performs in terms of robustness.

Nikoloulopoulos claims that the HK approximation can only be used if numbers of diseased and non-diseased are exactly equal across all studies. Otherwise, the underlying copula would not exist. However, this assertion comes without a proof, a reference, or at least a counterexample. For the HK model we assume beta-binomial distributions for the true positives (TP_i), the true negatives (TN_i) and the diseased probands of the i-th study $(i = 1, \dots I)$. This happens separately for each individual study, because we certainly wanted to adjust for different numbers of diseased and non-diseased probands across studies. As a consequence, we also assume a separate copula for each study whose existence is guaranteed by Sklar's theorem[5;6]. Each of these I copulas has the same association parameter which we estimate by standard maximum likelihood methods. Of course, the respective likelihood function includes the marginal beta parameters $(\alpha_{Se}, \beta_{Se}, \alpha_{Sp}, \beta_{Sp}, \alpha_{Pr}, \beta_{Pr})$ which are also assumed equal across studies. In essence, if there are relevant theoretical problems with the HK model, we would have expected to see serious problems in our simulations which deliberately were performed with differing numbers of diseased and non-diseased. The opposite was the case, even after varying the respective numbers of probands within studies between 20 and 300 we found satisfying results for sensitivity, specificity and prevalence.

Another advantage of the trivariate HK model is that it offers the possibility to account for the study design (case-control vs. cohort studies) as it was essential for the meta-analysis of the β-D-Glucan dataset[7] which was also used by Nikoloulopoulos. This is a simple consequence of the fact that disease prevalence in case-control studies is fixed

by design, and should not be considered as a random variable. As such, it becomes necessary to fit a bivariate model (neglecting prevalence) for the case-control studies, but simultaneously a trivariate model (including prevalence) for the cohort studies. In his example, Nikoloulopoulos only uses the cohort studies from the β-D-Glucan dataset and the question arises if his model also would be able to include case-control studies.

Finally, simulation studies should be designed to mimic real-world situations as perfect as possible. Only then the simulation results would be of help for the applied researcher who needs meta-analytic estimates for sensitivity and specificity for his data set at hand. For the HK paper we chose the true values for the different scenarios after analyzing a random sample of 15 meta-analyses compiled by Menke[8]. Therefore it would be helpful to know if Nikoloulopoulos has additional evidence that his simulation scenarios are more realistic than ours.

References

1. Nikoloulopoulos AK. A vine copula mixed effect model for trivariate meta-analysis of diagnostic test accuracy studies accounting for disease prevalence. *Statistical Methods in Medical Research* 2015; DOI: 10.1177/0962280215596769
2. Hoyer A and Kuss O. Statistical methods for meta-analysis of diagnostic tests accounting for prevalence - A new model using trivariate copulas. *Statistics in Medicine* 2015; **34**(11):1912-24. DOI: 10.1002/sim.6463
3. Nikoloulopoulos A. A mixed effect model for bivariate meta-analysis of diagnostic test accuracy studies using a copula representation of the random effects distribution. *Statistics in Medicine* 2015; DOI: 10.1002/sim.6595
4. Chu H, Nie L, Cole SR, Poole C. Meta-analysis of diagnostic accuracy studies accounting for disease prevalence: Alternative parameterizations and model selection. *Statistics in Medicine* 2009; **28**(18):2384-2399. DOI: 10.1002/sim.3627
5. Nelsen RB. *An Introduction to Copulas.* Springer: New York, 2006.
6. Sklar A. Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de L'Université de Paris* 1959; **8**:229-231.
7. Karageorgopoulos DE, Vouloumanou EK, Ntziora F, Michalopoulos A, Rafailidis PI, Falagas ME. β-D-Glucan Assay for the Diagnosis of Invasive Fungal Infections: A Meta-analysis. *Clinical Infectious Diseases* 2011; **52**(6):750-770. DOI: 10.1093/cid/ciq206
8. Menke J. Bivariate random-effects meta-analysis of sensitivity and specificity with SAS PROC GLIMMIX. *Methods of Information in Medicine* 2010; **49**(1):54-62, 62-64.