

Received XXXX

(www.interscience.wiley.com) DOI: 10.1002/sim.0000

Meta-analysis for diagnostic accuracy studies: A new statistical model using beta-binomial distributions and bivariate copulas

O. Kuss*, A. Hoyer, A. Solms

There are still challenges when meta-analysing data from studies on diagnostic accuracy. This is mainly due to the bivariate nature of the response where information on sensitivity and specificity must be summarized while accounting for their association within a single trial. In this paper we propose a new statistical model for the meta-analysis for diagnostic accuracy studies. This model uses beta-binomial distributions for the marginal numbers of true positives and true negatives and links these margins by a bivariate copula distribution. The new model comes with all the features of the current standard model, a bivariate logistic regression model with random effects, but has the additional advantages of a closed likelihood function, a larger flexibility for the association structure of sensitivity and specificity, and of operating on the original [0,1]-scale. In a simulation study, which compares three copula models and two implementations of the standard model, the Plackett copula model outperforms its competitors. An example from a meta-analysis to judge the diagnostic accuracy of telomerase (an urinary tumor marker) for the diagnosis of primary bladder cancer is used for illustration. Copyright © 2011 John Wiley & Sons, Ltd.

Keywords: Meta-analysis; Copula; Sensitivity; Specificity; Beta-binomial distribution

1. Introduction

While statistical methods for the meta-analysis of interventional trials are well-developed and understood nowadays, there are still challenges when meta-analysing data from studies on diagnostic accuracy. This is mainly due to the bivariate nature of the response, where information on sensitivity and specificity must be summarised while accounting for their association within a single trial. It is recommended [1, 2] that the bivariate logistic random-effects model (or the closely related [3] hierarchical summary ROC model) should be used for analysis. This model, when focussing on estimating meta-analytic sensitivities and specificities, has the advantages of accessing the individual data, and allowing unexplained heterogeneity as well as association between sensitivity and specificity [1]. Moreover, it can be generalized to model covariates, can cope with extreme values of 100% for sensitivity and specificity without applying artificial continuity corrections [4], and standard software (e.g., SAS[®], [5]) can be used for analysis. However, there are also some disadvantages of this standard model. It does not operate on the original scale of sensitivity and specificity, but on the corresponding logit scale, and, by generally relying on the bivariate normal distribution for the random effects, it only allows one single association structure. Most important, maximum likelihood estimation is complicated because the corresponding likelihood function has no closed form, but rather is a product of integrals which can not be solved analytically. This calls for numerical integration, MCMC techniques [6], or for approximative [5, 7] methods.

We propose here a new statistical model for the meta-analysis for diagnostic accuracy studies which avoids the previously mentioned problems while keeping all advantages of the standard model. It uses the idea of having marginal beta distributions for sensitivity and specificity, resulting in corresponding marginal beta-binomial distributions for true

Institute of Medical Epidemiology, Biostatistics, and Informatics, University of Halle-Wittenberg, Halle (Saale), Germany

* Correspondence to: Institut für Medizinische Epidemiologie, Biometrie und Informatik Medizinische Fakultät, Martin-Luther-Universität Halle-Wittenberg Magdeburger Str. 8, 06097 Halle (Saale)

positives and true negatives, and linking the marginals by copula distributions. This idea is not new for data like the one under consideration [8], but we feel that this is the first application of this idea in the medical sciences. Only recently, Chu et al. [9] used beta-binomial marginal distributions for the meta-analysis of bivariate response data but linked the marginal distributions by the Sarmanov family of bivariate distributions. Chen et al. [10] used the same approach in a Bayesian analysis for meta-analysis of case-control studies. This has the disadvantage that only a restricted range of values is allowed for the association parameters. Danaher/Smith [8] judged this model as ‘limited in its ability to model even moderate-sized levels of correlation’, which is especially problematic in the situation of diagnostic accuracy studies where large negative association values are the rule.

The paper is organized as follows. Section 2 introduces our model and points out the differences to the standard model. In section 3 we report on a simulation study that compares our model with the standard model. In section 4 we give an example, and section 5 discusses some additional technical points and concludes.

2. The Model

We begin by introducing the notation. Throughout we assume that each individual study (indexed by $i = (1, \dots, I)$) in the meta-analysis reports a four-fold table with the number of true positives (TP_i), true negatives (TN_i), false positives (FP_i), and false negatives (FN_i). The sensitivity in the i -th study (Se_i) is defined as $Se_i = TP_i / (TP_i + FN_i)$ and the specificity (Sp_i) as $Sp_i = TN_i / (TN_i + FP_i)$. The numbers of true positives and true negatives are assumed to be binomially distributed:

$$TP_i | Se_i \sim \text{Binomial}(TP_i + FN_i, Se_i), \quad TN_i | Sp_i \sim \text{Binomial}(TN_i + FP_i, Sp_i). \quad (1)$$

2.1. The standard model

In the standard model [4, 7] we further assume

$$\text{logit}(Se_i) = \mu + \phi_i, \quad \text{logit}(Sp_i) = \nu + \psi_i, \quad (2)$$

with

$$\begin{pmatrix} \phi_i \\ \psi_i \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\phi^2 & \rho\sigma_\phi\sigma_\psi \\ \rho\sigma_\phi\sigma_\psi & \sigma_\psi^2 \end{pmatrix} \right], \quad (3)$$

where $\text{logit}(p) = \log(p/(1-p))$ is the logit function. The parameters μ and ν are those of actual interest denoting the meta-analytic values for $\text{logit}(Se)$ and $\text{logit}(Sp)$, whereas the ϕ_i and ψ_i denote the normally distributed deviations from these global values for the i -th study. The association between Se_i and Sp_i is modelled via the parameter ρ in the random effects covariance matrix. This model constitutes a logistic regression model with random effects and a bivariate response. It thus belongs to the class of multivariate generalized linear mixed models, and the complete theory and software solutions for this model class apply. The number of parameters to be estimated equals five (μ , ν , and the three parameters of the random effects covariance matrix, ρ , σ_ϕ , and σ_ψ). Of course, the linear predictors $\text{logit}(Se_i)$ and $\text{logit}(Sp_i)$ can easily be extended to account for covariates or even additional random effects.

2.2. The new model

2.2.1. The marginal distributions Our model proceeds from the binomial distributions for TP_i and TN_i in (1) by making an assumption about Se and Sp on their original scale. To be concrete, we assume them to be Beta distributed with parameters a and b ,

$$Se \sim \text{Beta}(a_{Se}, b_{Se}), \quad Sp \sim \text{Beta}(a_{Sp}, b_{Sp}), \quad (4)$$

and density

$$f(p; a_{Se}, b_{Se}) = P(p = u) = \frac{\Gamma(a_{Se} + b_{Se})}{\Gamma(a_{Se})\Gamma(b_{Se})} u^{a_{Se}-1} (1-u)^{b_{Se}-1} \quad (5)$$

for Se , and analogously defined for Sp . The expectations ($a_*/(a_* + b_*)$) of this two beta distributions describe the meta-analytic parameters of interest (Se , Sp), the variances ($a_*b_*/(a_* + b_* + 1)(a_* + b_*) * 2$) their variations which correspond to heterogeneity in the meta-analytic sense.

As the Beta distribution is conjugate to the binomial distribution we achieve a Beta-binomial distribution for TP_i and TN_i with density:

$$P(TP_i = k) = \binom{TP_i + FN_i}{k} \frac{B(a_{Se} + k, TP_i + FN_i - k + b_{Se})}{B(a_{Se}, b_{Se})}, \quad (6)$$

where B is Euler's Beta function, and analogously defined for TN_i . It should be noted that this Beta-binomial model is a true random effects model in the sense that each subject (here: each single study) has a single parameter (here: Se_i or Sp_i) where this parameters follow a joint common distribution. However, and in contrast to most other random effects models, the resulting marginal distribution has still closed form [11].

2.2.2. Copulas For modelling the association between TP_i and TN_i we apply the concept of copulas. The theory of copulas is a mighty statistical theory on its own and Nelsen [12] gives a comprehensive account of it. For the task here it suffices to consider only bivariate copulas, and we use the more accessible description of Danaher/Smith [8]. Consider two random variables X_1 and X_2 with distribution functions $F(X_1)$ and $F(X_2)$. Sklar [13] proved that there exists a function C with

$$H(X_1 = x_1, X_2 = x_2) = C(F(x_1), F(x_2)) = C(u_1, u_2), \quad (7)$$

where $C(u_1, u_2)$ is itself a distribution function for a bivariate pair for uniform random variables and $H(X_1 = x_1, X_2 = x_2)$ is a distribution function for the original variables X_1 and X_2 . If $C(u_1, u_2)$ fulfils three rather unrestrictive properties [12], it is called a 'copula' and its exclusive role is to determine the dependence between $F(X_1)$ and $F(X_2)$ and thus between X_1 and X_2 . In particular, this is done by using only the original marginals, keeping these completely independent from the copula association parameter.

By using copulas we are able to reduce our model for the meta-analysis of diagnostic accuracy studies to simply fitting a bivariate distribution. This is done by differentiating the copula, interpreting the resulting density as the likelihood function, and use standard maximum likelihood methods for parameter estimation. It is especially convenient that this likelihood has closed form. Again, and as compared to the standard model, there are five model parameters: Two for the beta-binomial distribution of each margin, and one additional copula parameter that controls the association between the margins. We mention one technical issue which we appreciate in full detail in the discussion: The straightforward differentiation of the copula with respect to the Lebesgue measure requires the marginals to be continuous. This is not the case here, beta-binomial distributions are discrete ones which calls for differentiation with respect to the counting measure.

To quantify the association between X_1 and X_2 , it was shown [14] that both, Spearman's $\rho_S(X_1, X_2)$ and Kendall's $(\tau(X_1, X_2))$ correlation coefficients can be described solely in terms of the copula by

$$\rho_S(X_1, X_2) = 12E\{(F_1(x_1) - 1/2)(F_2(x_2) - 1/2)\} = 12 \iint \{C(u, v) - uv\} dudv \quad (8)$$

and

$$\tau(X_1, X_2) = P\{(X_1 - X_1^*)(X_2 - X_2^*) > 0\} - P\{(X_1 - X_1^*)(X_2 - X_2^*) < 0\} = 4 \iint C(u, v) dC(u, v) - 1, \quad (9)$$

where (X_1^*, X_2^*) is an independent copy of (X_1, X_2) to define $\tau(X_1, X_2)$ as a measure of concordance. It should be noted that the Pearson correlation of X_1 and X_2 in general depends also on the marginal distributions and is thus affected by scale changes.

It is immediate and an advantage of the new model that a rich number of copulas could be used, each one resulting in a new model for the meta-analysis of diagnostic accuracy studies. This should be contrasted to the standard model that, by referring to the bivariate normal distribution, only allows one single association structure. In practice however, this set of copulas is reduced by the specification that it should be possible to model the whole range of associations from perfectly negative to perfectly positive. To be specific, we use three copulas here, the Clayton [15], the Gauss, and the Plackett copula, all of them allowing the full range of associations.

Clayton copula The Clayton copula [15] is defined the following way:

$$C_C(u_1, u_2, \theta_C) = (\max(u_1^{-\theta_C} + u_2^{-\theta_C} - 1, 0))^{-\frac{1}{\theta_C}}, \theta_C \in [-1, \infty) \setminus \{0\}. \quad (10)$$

Its density is given by

$$c_C(u_1, u_2, \theta_C) = (1 + \theta_C)(u_1 u_2)^{-\theta_C - 1} (u_1^{-\theta_C} + u_2^{-\theta_C} - 1)^{-2 - \frac{1}{\theta_C}} \quad (11)$$

in the continuous case. The Clayton copula is a member of the class of Archimedean copulas. For measuring association between the marginals we can use Kendall's Tau with $\tau = \theta_C / (\theta_C + 2)$.

Gauss copula The Gauss copula is defined by

$$C_G(u_1, u_2, \Gamma) = \Phi_2(\Phi^{-1}(u_1), \Phi^{-1}(u_2)|\Gamma) \quad (12)$$

with $\Phi_2(\cdot|\Gamma)$ as the distribution function of a bivariate standard normal distribution $N(0, \Gamma)$ with correlation matrix Γ which in this case is equal to its covariance matrix. Φ^{-1} is the distribution function of the standard (univariate) normal distribution. The density of the Gauss copula can be shown to be [16]

$$c_G(u_1, u_2, \Gamma) = |\Gamma|^{-1/2} \exp\left\{\frac{1}{2} \mathbf{q}^T (I_2 - \Gamma^{-1}) \mathbf{q}\right\} \quad (13)$$

with $\mathbf{q} = (q_1, q_2)^T$ the vector of normal scores, that is $q_j = \Phi^{-1}(u_j)$, $j = 1, 2$, and I_2 the two-dimensional identity matrix. The Gauss copula is a member of the class of elliptical copulas which have the advantage that they can easily be generalized to more than two margins. Interpretation of the association parameter γ (the off-diagonal element of Γ) is not straightforward. Song [16] shows that it equals the Pearson correlation of the two normal scores q_1 and q_2 , which is in general numerically close to the Spearman correlation of the original margins.

Plackett copula The Plackett copula originates from a family of bivariate distributions with uniform margins that was proposed by Plackett [17]. Later Nelsen [12] showed that the Plackett family can also be interpreted as a copula with the original index of the family measuring the association between the margins. The Plackett copula is defined via

$$C_P(u_1, u_2, \theta_P) = \frac{(1 + (\theta_P - 1)(u_1 + u_2)) - \sqrt{(1 + (\theta_P - 1)(u_1 + u_2))^2 - 4u_1u_2\theta_P(\theta_P - 1)}}{2(\theta_P - 1)}, \theta_P > 0, \quad (14)$$

and the density is

$$c_P(u_1, u_2, \theta_P) = ((1 + (\theta_P - 1)(u_1 + u_2))^2 - 4\theta_P(\theta_P - 1)u_1u_2)^{-\frac{3}{2}} \theta_P (1 + (\theta_P - 1)(u_1 + u_2 - 2u_1u_2)). \quad (15)$$

Sperman's ρ equals

$$\rho_S = \frac{\theta_P + 1}{\theta_P - 1} - \frac{2\theta_P}{(\theta_P - 1)^2} \log(\theta_P). \quad (16)$$

Nelsen [12] showed that the parameter θ_P from the Plackett copula could also be interpreted as an odds ratio from a fourfold table which arises from a dichotomization of the marginals at points x_1 and x_2 , where this odds ratio is constant for all pairs of (x_1, x_2) .

3. Simulation

To compare the statistical properties of our copula models to the standard model we conducted a simulation study. We generated meta-analyses of diagnostic accuracy studies from four different scenarios (standard model, and each of the three copula introduced above) and estimated sensitivity and specificity with five different methods, two from the standard model (Gaussian quadrature with SAS NLMIXED and penalized quasi-likelihood with SAS GLIMMIX) and one for each copula model where we used maximum likelihood estimation with SAS NLMIXED. We further varied

- the number of studies in the respective meta-analysis (10 or 50),
- the true sensitivity and specificity (70%/70%, 90%/70%, or 90%/90%),
- the true variance of sensitivity and specificity (small or large),
- the true association between sensitivity and specificity (none, small or large), and
- the number of patients per study (20 or 100).

With the true model being one of the copula models, a small/large variance corresponded to a value of 0.01/0.05 on the [0,1] scale. When the true model was the standard model, these values were 0.15/0.75 on the logit scale. In terms of associations, a small/large association stood for a correlation of approximately -0.2/-0.8 on the respective scale. True values for the different scenarios were chosen a priori and after analysing a sample of 15 meta-analyses compiled by Menke [5]. Combination of the six design factors resulted in a total of 288 simulation scenarios. For each scenario, 1000 meta-analyses were generated. Following the current guideline of Burton et al. [18] we used bias, mean squared error (MSE) and empirical coverage (to the 95%-level) as outcomes to compare the five different methods. To reduce complexity, we averaged biases, mean squared errors and coverages for sensitivity and specificity from each meta-analysis. To assess numerical robustness of the different procedures, we also report the number of converged simulation runs. The simulation

programme was written in SAS. To enable fair comparison between models, all SAS procedures for parameter estimation were run with the default options. Starting values (which must be given for the NLMIXED procedure) were computed as raw proportions for sensitivity and specificity with corresponding binomial variances for each single meta-analysis separately. The starting values for the association parameters were estimated from the raw correlation and transformed as appropriate for the respective model.

For the description of results we draw boxplots for the distributions of the four outcomes (bias, MSE, coverage, and number of converged simulation runs) across the simulation scenarios. In 37 scenarios the SAS code broke down completely due to floating point errors in one of the copula models and as these scenarios are expected to be the most numerically challenging we used only the 251 simulation scenarios which yielded complete results for all 5 models. To assess the influence of the different factors we computed, for each outcome separately, linear regression models with the six factors as covariates and report least-square means.

Referring to the results, the Gauss (CG) and the Plackett (CP) copula were superior in terms of bias (Figure 1), whereas the Plackett copula (CP) and the PQL estimation from SAS GLIMMIX (SG) were superior in terms of mean squared error (Figure 2). Regarding coverage, none of the procedures gave satisfactory results, all missed the designated 95% coverage considerably (Figure 3). However, the Plackett copula (CG) and SAS GLIMMIX (SG) still outperform the competing models in terms of coverage. Referring to the number of converged simulation runs, also the Gauss (CG) and the Plackett (CP) copula were superior to the other procedures, however, it has to be remembered that there were 31 simulation scenarios where one of the copula codes broke down completely. Not surprisingly (table 1), outcomes were improved by a larger number of studies, a weaker association between sensitivity and specificity, smaller variances of sensitivity and specificity, and larger numbers of patients with studies. No clear influence could be seen across the different patterns for the values of sensitivity and specificity.

Summing up the results from the simulation study, the Plackett copula was the superior copula model, and PQL estimation outperformed Gaussian quadrature (SN) in the group of the standard models. Comparing copula and standard models, the Plackett copula performed best.

PLACE FIGURES 1-4 AND TABLE 1 APPROXIMATELY HERE

4. Example

As an example we use the meta-analysis of Glas et al. [19] to judge the diagnostic accuracy of telomerase (an urinary tumor marker) for the diagnosis of primary bladder cancer, where it was of interest if this non-invasive and cheap marker could replace the then standard of cystoscopy and/or histopathology. The data set (see table 2) was used at several instances in the methodical literature [7, 20], mainly because the standard model had problems to give sensible parameter estimates. This was attributed to the large negative value of the association between sensitivity and specificity. Paul et al. [7] and Riley et al. [20] reported results for the standard model with SAS NLMIXED, however, only after trying a range of starting values. Glas et al. [19] in the original analysis applied a standard two-step approach (estimate sensitivities and specificities and their standard errors in a first step, combine them with respective weights in the second step) to achieve their parameter estimates. In the appendix we give the SAS and R code to analyse the telomerase data by the Plackett copula model.

PLACE TABLE 2 APPROXIMATELY HERE

In table 3 we give the results for the telomerase data for our three copula models and for the standard models. Where most models roughly agree on the estimated sensitivity there are some differences between estimated specificities which range from 81.9% from the Gauss copula to 91.2% from the standard model with Gaussian quadrature. The Plackett copula which performed best in our simulation study finds a specificity of 86.5%. It should be noted that the Gauss and the Plackett copula yield sensible estimates for the association parameters which are additionally accompanied by a 95% confidence interval.

PLACE TABLE 3 APPROXIMATELY HERE

5. Discussion

In this paper we proposed a new statistical model for the meta-analysis of diagnostic accuracy studies. This model uses beta-binomial distributions for the marginal numbers of true positives and true negatives and links these marginals by a copula distribution. This model comes with all the features of the current standard model, a bivariate logistic regression

model with random effects, but has the additional advantages of a closed likelihood function, a larger flexibility for the association structure of sensitivity and specificity, and of operating on the original $[0,1]$ -scale. As compared to a recent implementation of an approximate MCMC method [7] for the standard model there is also no need for the specification of prior distributions. In a simulation study, the Plackett copula model outperformed the standard model.

It is fair to discuss some limitations of our proposed model. First, the model constitutes an actual random effects model in the meta-analytic sense. That is, sensitivity and specificity are assumed to vary randomly around an underlying average value and the estimated variances will always be positive and larger than zero. That is, the model will not simplify to a fixed effects model in the case of homogeneity as a standard meta-analytic model does. However, if one wishes to model one marginal as fixed, one could skip the beta-binomial distribution and use the standard binomial distribution for the respective margin. The copula idea would still apply here because it is not necessary that all marginals are from the same family of distributions. Moreover, and as pointed out by Chu et al. [21], there is a great potential for heterogeneity in the meta-analysis for diagnostic accuracy studies, where this heterogeneity between studies arises due to differences in disease prevalence, study population characteristics or laboratory methods. Second, it is justified to ask if the beta distribution that is assumed for the distribution of sensitivity and specificity across studies in our model is flexible enough, especially if compared to the respective class from the standard model. Aitchison/Begg [22] termed the corresponding class of distributions from the standard model 'logistic-normal distributions' and showed that these indeed yield slightly more flexible models. However, both classes are restricted to 'U-shaped, J-shaped, and flattish to sharpish unimodal curves.' Third, although it is a definite advantage of copulas that they model the association between variables strictly separated from the marginal distributions of these variables, it is obvious that the estimation of marginals and association interferes if both are collected in a common likelihood function. To be specific, if we assume the wrong copula for modelling the association structure we might also get compromised estimates for the parameters of the marginal distributions. However, this problem is also a threat for the standard model which relies on a bivariate normal distribution on the logit-scale. Fourth, despite the fact that the Plackett copula outperformed both estimation methods for the standard model, we still might not be satisfied with its performance in terms of coverage. The observed median value of 86.6% across all simulation settings is still way below the designated 95%. This deficiency is most probably caused by the usage of the standard t confidence interval here, and further work should look for a effective correction of the respective degrees of freedom, or maybe turn to profile likelihood confidence intervals. For example, profile likelihood confidence intervals for the telomerase data (calculated by SAS NLP) from the Plackett copula were (72.9%, 82.1%) for the sensitivity and (70.1%, 93.8%) for the specificity. While the interval lengths are comparable, the confidence interval for specificity is considerably shifted to the left. Fifth, the copula models are not suited to deal with studies that report only either a sensitivity or a specificity, where this deficiency is also given for the standard model. In this case, we might turn to a multiple imputation technique to replace the missing value.

In the introduction we mentioned the problem that the differentiation of the copula which is needed to achieve the likelihood function for our model should be performed with respect to the Lebesgue measure. This is in conflict to our marginal beta-binomial distributions which are discrete and should be differentiated with respect to the counting measure. Genest/Nešlehová [23] pointed to some of the problems of copulas with count data, however, their examples are extreme ones with very small numbers of support points for the marginal distributions. Luckily, the likelihood functions for our model can also be derived with respect to the counting measure and they are only slightly more complicated than those for continuous marginals. We assessed the discrete likelihoods in a sub-sample of 1000 meta-analyses from our whole simulation sample and found (data not shown) no superior performance for the discrete likelihoods. Instead, using the discrete likelihoods resulted in a considerably larger number of non-converged models.

Our model could be generalized in a number of directions. For example, we might enhance the model in the sense of Chu et al. [21] by additionally accounting for disease prevalence. This disease prevalence could be incorporated as a third marginal distribution, and sensitivity, specificity, and disease prevalence could be linked together by a trivariate copula. Trivariate copula can be straightforwardly defined and estimated in the elliptical case (that is, for a generalization of the Gaussian copula), but there also exists a trivariate version of the Plackett copula [24]. Moreover, we could go for different marginal distributions or use two-parameter copulas. Finally, we could use the idea of beta-binomial marginals and copula distributions also for other areas of meta-analysis, e.g. for the meta-analysis of intervention trials.

References

1. Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM; Cochrane Diagnostic Test Accuracy Working Group. Systematic reviews of diagnostic test accuracy. *Annals of Internal Medicine*. 2008; 149(12):889-97.
2. Harbord RM, Whiting P, Sterne JA, Egger M, Deeks JJ, Shang A, Bachmann LM. An empirical comparison of methods for meta-analysis of diagnostic accuracy showed hierarchical models are necessary. *Journal of Clinical Epidemiology*. 2008; 61(11):1095-103.
3. Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JA. A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics*. 2007; 8(2):239-51.
4. Chu H, Cole SR. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. *Journal of Clinical Epidemiology*. 2006; 59(12):1331-2.
5. 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, 62-4.
6. Verde PE. Meta-analysis of diagnostic test data: a bivariate Bayesian modeling approach. *Statistics in Medicine*. 2010; 29(30):3088-102.
7. Paul M, Riebler A, Bachmann LM, Rue H, Held L. Bayesian bivariate meta-analysis of diagnostic test studies using integrated nested Laplace approximations. *Statistics in Medicine*. 2010; 29(12):1325-39.
8. Danaher PJ, Smith MS. Modeling Multivariate Distributions Using Copulas: Applications in Marketing. *Marketing Science*. 2011; 30(1):4-21.
9. Chu H, Nie L, Chen Y, Huang Y, Sun W. Bivariate random effects models for meta-analysis of comparative studies with binary outcomes: Methods for the absolute risk difference and relative risk. *Statistical Methods in Medical Research*. 2010 Dec 21. [Epub ahead of print]
10. Chen Y, Chu H, Luo S, Nie L, Chen S. Bayesian analysis on meta-analysis of case-control studies accounting for within-study correlation. *Statistical Methods in Medical Research*. 2012
11. Molenberghs G, Verbeke G. *Models for Discrete Longitudinal Data*. Springer: New York, 2006.
12. Nelsen RB. *An Introduction to Copulas*. Springer: New York, 2006.
13. Sklar A. Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de L'Université de Paris*. 1959; 8:229231.
14. Schweizer B, Wolff EF. On nonparametric measures of dependence for random variables. *Annals of Statistics*. 1981; 9:879-885.
15. Clayton DG. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*. 1978; 65(1):141-151.
16. Song PXX. Multivariate Dispersion Models Generated From Gaussian Copula. *Scandinavian Journal of Statistics*. 2000; 27(2):305-320.
17. Plackett RL. A Class of Bivariate Distributions. *Journal of the American Statistical Association*. 1965; 60(310):516-522.
18. Burton A, Altman DG, Royston P, Holder RL. The design of simulation studies in medical statistics. *Statistics in Medicine*. 2006;25(24):4279-92.
19. Glas AS, Roos D, Deutekom M, Zwinderman AH, Bossuyt PM, Kurth KH. Tumor markers in the diagnosis of primary bladder cancer. A systematic review. *Journal of Urology*. 2003; 169(6):1975-82.
20. Riley RD, Abrams KR, Sutton AJ, Lambert PC, Thompson JR. Bivariate random-effects meta-analysis and the estimation of between-study correlation. *BMC Medical Research Methodology*. 2007;7:3.
21. 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):2384-99.
22. Aitchison J, Begg CB. Statistical Diagnosis When Basic Cases are Not Classified with Certainty. *Biometrika*. 1976; 63(1), 1-12.
23. Genest C, Nešlehová J. A Primer on Copulas for Count Data. *Astin Bulletin*. 2007; 37(2), 475-515.
24. Kao SC, Govindaraju RS. Trivariate statistical analysis of extreme rainfall events via the Plackett family of copulas. *Water Resources Research*. 2008; 44:W02415.

6. Appendix

In this appendix we give the SAS and R code to fit the Plackett copula model for the telomerase data. Starting values for the marginal distributions were computed from raw marginal expectations and variances, the starting value for the association parameter was computed from the raw estimate of Kendall's tau.

6.1. SAS NLMIXED

```

DATA telomerase;
  INPUT tp fn tn fp @@;
  s=fn+tp;h=fp+tn;
  CARDS;
  25 8 25 1 17 4 11 3 88 16 31 16 16 10 80 3 40 17 137 1
  38 9 24 6 23 19 12 0 27 6 18 2 14 3 29 3 37 7 7 22
;RUN;

PROC NLMIXED DATA=telomerase ;
  * Give the parameters to be estimated along with their starting values;
  PARSMS param_se=1.136 param_sp=1.524 rho_se=0.011 rho_sp=0.053 temptheta=0.502;

  * Transform the binomial probabilities to the expit-scale to avoid boundary problems;
  pi_se = exp(param_se)/(1 + exp(param_se));
  pi_sp = exp(param_sp)/(1 + exp(param_sp));

  * Link the transformed binomial probabilities with the parameters from the beta distributions;
  a_se = pi_se *(1-rho_se)/rho_se ;
  b_se = (1-pi_se)*(1-rho_se)/rho_se ;
  a_sp = pi_sp *(1-rho_sp)/rho_sp ;
  b_sp = (1-pi_sp)*(1-rho_sp)/rho_sp ;

  * Beta-binomial loglikelihood functions for sensitivity and specificity;
  ll_se = LGAMMA (s+1)+LGAMMA (tp+a_se)+LGAMMA (s-tp+b_se)+LGAMMA (a_se+b_se)
        -LGAMMA (tp+1)-LGAMMA (s-tp+1)-LGAMMA (s+a_se+b_se)-LGAMMA (a_se)-LGAMMA (b_se) ;
  ll_sp = LGAMMA (h+1)+LGAMMA (tn+a_sp)+LGAMMA (h-tn+b_sp)+LGAMMA (a_sp+b_sp)
        -LGAMMA (tn+1)-LGAMMA (h-tn+1)-LGAMMA (h+a_sp+b_sp)-LGAMMA (a_sp)-LGAMMA (b_sp) ;

  * Define the distribution functions of the beta-binomial distribution from scratch
  as they are not implemented as SAS functions;
  F_se = 0;
  DO i = 0 TO tp;
    f_se_pdf = exp (LGAMMA (s+1)+LGAMMA (i+a_se)+LGAMMA (s-i+b_se)+LGAMMA (a_se+b_se)
                  -LGAMMA (i+1)-LGAMMA (s-i+1)-LGAMMA (s+a_se+b_se)-LGAMMA (a_se)-LGAMMA (b_se) ) ;
    F_se = F_se + f_se_pdf;
  END;
  F_sp = 0;
  DO i = 0 TO tn;
    f_sp_pdf = exp (LGAMMA (h+1)+LGAMMA (i+a_sp)+LGAMMA (h-i+b_sp)+LGAMMA (a_sp+b_sp)
                  -LGAMMA (i+1)-LGAMMA (h-i+1)-LGAMMA (h+a_sp+b_sp)-LGAMMA (a_sp)-LGAMMA (b_sp) ) ;
    F_sp = F_sp + f_sp_pdf;
  END;

  * Define the Plackett density;
  theta = exp(temptheta);
  ll_plackett = -3/2*log((1 + (theta - 1)*(F_se + F_sp))**2 - 4*theta*(theta - 1)*F_se*F_sp)
               + log(theta*(1 + (theta - 1)*(F_se + F_sp- 2*F_se*F_sp)));

  * Collect the elements of the loglikelihood function;
  ll = ll_se + ll_sp + ll_plackett;

  MODEL tp ~ GENERAL(ll);

  * Estimate the parameters of interest;
  ESTIMATE "Sensitivity" exp(param_se)/(1 + exp(param_se)) ;
  ESTIMATE "Specificity" exp(param_sp)/(1 + exp(param_sp)) ;
  ESTIMATE "Plackett's Tau" (theta+1)/(theta-1)-((2*theta)/(theta-1)**2)*log(theta);
RUN;

```

6.2. R

```

# Load packages
library(VGAM)
library(copula)
library(ucminf)

telomerase = data.frame(rbind(
  c(25, 1, 8, 25), # 1. Ito et al.
  c(17, 3, 4, 11), # 2. Rahat et al.
  c(88, 16, 16, 31), # 3. Kavalier et al.
  c(16, 3, 10, 80), # 4. Yoshida et al.
  c(40, 1, 17, 137), # 5. Ramakumar et al.
  c(38, 6, 9, 24), # 6. Landman et al.
  c(23, 0, 19, 12), # 7. Kinoshita et al.
  c(27, 2, 6, 18), # 8. Gelmini et al.
  c(14, 3, 3, 29), # 9. Cheng et al.
  c(37, 22, 7, 7) # 10. Cassel et al.
))

names(telomerase) = c("TP", "FP", "FN", "TN")

# Define bivariate beta binomial log-likelihood function
# with Plackett copula

logLL.BivBetaBin = function(data, parms)
{
  # data: observations with structure dataframe or matrix
  # colnames = "TP", "FP", "FN", "TN"
  # parms: named parameter vector,
  # names = "se", "sp", "rho_se", "rho_sp", "ltheta"

  # Number of studies
  N = length(data[,1])

  .se = parms["se"]; .rho_se = parms["rho_se"]
  .sp = parms["sp"]; .rho_sp = parms["rho_sp"]

  # Transform the binomial probabilities to the expit-scale to
  # avoid boundary problems
  rho_se = exp(.rho_se)/(1+exp(.rho_se))
  rho_sp = exp(.rho_sp)/(1+exp(.rho_sp))
  pi_se = exp(.se)/(1+exp(.se))
  pi_sp = exp(.sp)/(1+exp(.sp))

  # Define the Plackett copula
  cop = plackettCopula(param = exp(parms["ltheta"]))

  fn = data[,"FN"]; tp = data[,"TP"]; s = fn+tp
  fp = data[,"FP"]; tn = data[,"TN"]; h = fp+tn

  # Beta-binomial log-likelihood functions for sensitivity and specificity
  ll_se = dbetabinom(x = tp, size = s, prob = pi_se, rho = rho_se, log = TRUE)
  ll_sp = dbetabinom(x = tn, size = h, prob = pi_sp, rho = rho_sp, log = TRUE)

  # Beta-binomial probability distribution function & log-likelihood of the copula
  F_se = vector(); F_sp = vector(); ll_copula = vector()
  for (i in 1:N)
  {
    F_se[i] = pbetabinom(q = tp[i], size = s[i], prob = pi_se, rho = rho_se)
    F_sp[i] = pbetabinom(q = tn[i], size = h[i], prob = pi_sp, rho = rho_sp)
    ll_copula[i] = log(dcopula(cop, c(F_se[i],F_sp[i])))
  }

  # Collect the elements of the loglikelihood function
  logLL = ll_se + ll_sp + ll_copula
}

```

```

-sum(logLL)
}

# Calculate standard errors and confidence intervals by the multivariate delta method

results = function(est, conf.level, df)
{
# 1.) est = output of ucminf fitting procedure
# 2.) conf.level = 0.95 for 95% Confidence Intervals
# 3.) df = degree of freedom for students-t distribution (here numbers of studies)

# Original Parameters
.sens = exp(est$par["se"])/(1+exp(est$par["se"]))
.spec = exp(est$par["sp"])/(1+exp(est$par["sp"]))
.theta = exp(est$par["ltheta"])

# Plackett' tau
.tau = (.theta+1)/(.theta-1)-2*.theta/((.theta-1)**2*log(.theta))

# Inverse Fisher information matrix
.invFish = est$invhessian

# Standard errors of transformed parameters
.SE = sqrt(diag(.invFish))
names(.SE) = names(est$par)

# Standard Errors of original parameters
.SE["sens"] = .sens*(1-.sens)*.SE["se"]
.SE["spec"] = .spec*(1-.spec)*.SE["sp"]
.SE["theta"] = .theta*.SE["ltheta"]
.SE["tau"] = -2/((.theta-1)**2*(2-log(.theta)*(.theta+1)/(.theta-1)))*.SE["theta"]

# Criterion for confidence intervals
.crit = qt(1-(1-conf.level)/2, df = df-1)

.CIsens = .sens+c(-1,1)*.crit*.SE["sens"]
.CIspec = .spec+c(-1,1)*.crit*.SE["spec"]
.CItau = .tau +c(-1,1)*.crit*.SE["tau"]

out = data.frame(rbind(
  par = c(.sens, .spec, .tau),
  SE = c(.SE["sens"], .SE["spec"], .SE["tau"]),
  CI_lb = c(.CIsens[1], .CIspec[1], .CItau[1]),
  CI_up = c(.CIsens[2], .CIspec[2], .CItau[2])))
names(out) = c("sens", "spec", "tau")
out
}

# Give the parameters along with their starting values
initial = c(se = 1.136, rho_se = 0.011,
  sp = 1.524, rho_sp = 0.053,
  ltheta = 0.502)

# Fit the data
fit = ucminf(par = initial,
  fn = logLL.BivBetaBin,
  data = telomerase,
  hessian = 3)
(results(est=fit, conf.level=0.95, df=10))

```

Table 1. Results from the linear regression models to assess the influence of the different factors in the simulation study. Given are the least-square means with their 95%-confidence intervals. The following abbreviations are used: CC=Clayton copula, CG=Gauss copula, CP=Plackett copula, SG=Standard model estimated by penalized quasi-likelihood in SAS GLIMMIX, and SN=Standard model estimated by Gaussian quadrature in SAS NLMIXED. In terms of the true model "SM" denotes meta-analyses generated from the standard model. The estimates for the mean squared error were computed with values above 10 in the SN-model truncated to 10.

Factor	Value	Bias (95%-CI)		MSE (95%-CI)		Coverage (95%-CI)	
Estimated model	CC	3.34	(2.95, 3.74)	0.69	(0.48, 0.91)	50.1	(47.2, 53.1)
	CG	-0.78	(-1.18, -0.39)	0.62	(0.41, 0.84)	68.3	(65.4, 71.3)
	CP	0.52	(0.13, 0.91)	0.29	(0.07, 0.50)	74.4	(71.5, 77.4)
	SG	2.57	(2.18, 2.96)	0.44	(0.23, 0.66)	75.1	(72.1, 78.0)
	SN	3.24	(2.84, 3.63)	2.23	(2.02, 2.45)	59.0	(56.1, 61.9)
True model	CC	2.32	(1.96, 2.68)	0.88	(0.68, 1.08)	64.9	(62.1, 67.6)
	CG	2.08	(1.71, 2.44)	1.00	(0.82, 1.18)	67.8	(65.1, 70.5)
	CP	3.23	(2.86, 3.60)	0.66	(0.46, 0.86)	60.3	(57.5, 63.0)
	SM	-0.51	(-0.84, -0.19)	0.88	(0.68, 1.08)	68.7	(66.2, 71.1)
Number of studies	10	1.70	(1.42, 1.98)	0.57	(0.44, 0.70)	72.6	(70.5, 74.7)
	50	1.86	(1.63, 2.09)	1.14	(0.99, 1.30)	58.2	(56.4, 59.9)
True association between sensitivity and specificity	strong	3.22	(2.91, 3.53)	1.12	(0.95, 1.29)	54.0	(51.7, 56.3)
	weak	1.52	(1.21, 1.82)	0.59	(0.42, 0.75)	68.3	(66.0, 70.6)
	none	0.59	(0.29, 0.90)	0.86	(0.69, 1.03)	73.9	(71.6, 76.2)
True sensitivity and specificity	90%/90%	0.98	(0.66, 1.31)	0.89	(0.72, 1.07)	58.2	(55.8, 60.7)
	90%/70%	2.27	(1.95, 2.59)	0.69	(0.52, 0.87)	65.1	(62.7, 67.5)
	70%/70%	2.08	(1.80, 2.36)	0.98	(0.82, 1.14)	72.8	(70.7, 74.9)
True Variance	Large	2.23	(1.95, 2.51)	0.91	(0.76, 1.06)	65.5	(63.4, 67.6)
	Small	1.32	(1.09, 1.55)	0.80	(0.67, 0.93)	65.3	(63.6, 67.0)
Number of patients per study	20	1.72	(1.47, 1.98)	0.78	(0.64, 0.92)	59.8	(57.9, 61.7)
	100	1.83	(1.58, 2.08)	0.93	(0.79, 1.07)	71.0	(69.1, 72.8)

Table 2. Example data set for the telomerase data from Glas et al. [19]

Study	TP	FP	FN	TN
Itô et al.	25	1	8	25
Rahat et al.	17	3	4	11
Kavaler et al.	88	16	16	31
Yoshida et al.	16	3	10	80
Ramakumar et al.	40	1	17	137
Landman et al.	38	6	9	24
Kinoshita et al.	23	0	19	12
Gelmini et al.	27	2	6	18
Cheng et al.	14	3	3	29
Cassel et al.	37	22	7	7

Table 3. Results for the different estimation methods for the telomerase data from Glas et al. [19].

Method	Sensitivity (95%-CI)	Specificity (95%-CI)	Association (95%-CI)
Clayton	–	–	–
Gauss	79.4% (75.1%, 83.7%)	81.9% (69.0%, 94.7%)	$\gamma = -0.71$ (-0.95, -0.47)
Plackett	77.6% (72.6%, 82.6%)	86.5% (74.7%, 98.3%)	$\tau = -0.85$ (-1.03, -0.68)
Standard model (GQ, NLMIXED)	76.7% (69.4%, 84.0%)	91.2% (79.6%, 100.3%)	$\rho = -1.00$ (-, -)
Standard model (PQL, GLIMMIX)	76.7% (68.8%, 83.1%)	90.4% (71.5%, 97.2%)	$\rho = -1.00$ (-, -0.33)
Standard model (MCMC) [7]	76.7% (69.8%, 82.6%)	90.9% (73.7%, 97.8%)	$\rho = -0.88$ (-0.99, -0.18)
Original analysis [19]	75% (71%, 79%)	86% (71%, 94%)	$\rho' = -0.73$ (-, -)

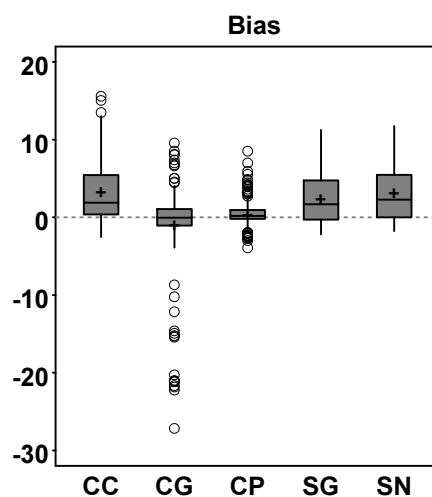


Figure 1. Distribution of bias across the 251 simulation scenarios which gave complete results for all five procedures, given as boxplots. The following abbreviations are used: CC=Clayton copula, CG=Gauss copula, CP=Plackett copula, SG=Standard model estimated by penalized quasi-likelihood in SAS GLIMMIX, and SN=Standard model estimated by Gaussian quadrature in SAS NLMIXED.

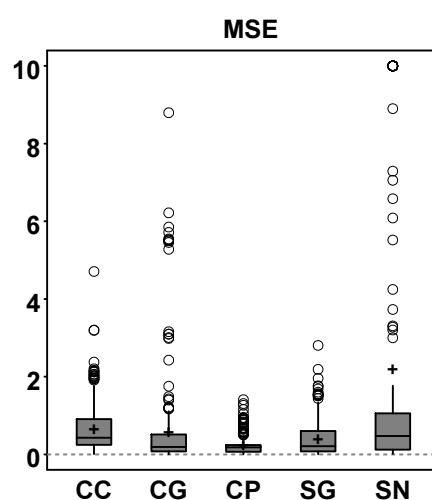


Figure 2. Distribution of mean squared error (MSE) across the 251 simulation scenarios which gave complete results for all five procedures, given as boxplots. To enhance readability of the plot, all values larger than 10 were set to 10. The following abbreviations are used: CC=Clayton copula, CG=Gauss copula, CP=Plackett copula, SG=Standard model estimated by penalized quasi-likelihood in SAS GLIMMIX, and SN=Standard model estimated by Gaussian quadrature in SAS NLMIXED.

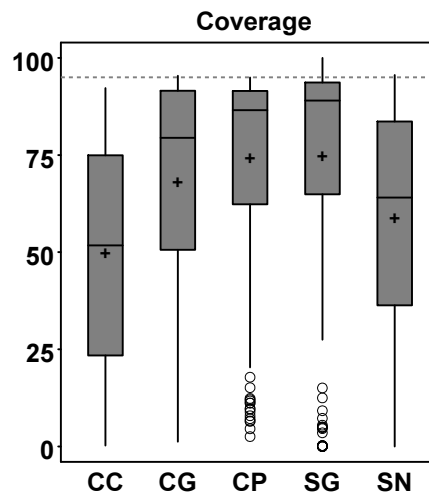


Figure 3. Distribution of coverage to the 95% level across the 251 simulation scenarios which gave complete results for all five procedures, given as boxplots. The following abbreviations are used: CC=Clayton copula, CG=Gauss copula, CP=Plackett copula, SG=Standard model estimated by penalized quasi-likelihood in SAS GLIMMIX, and SN=Standard model estimated by Gaussian quadrature in PROC NLMIXED.

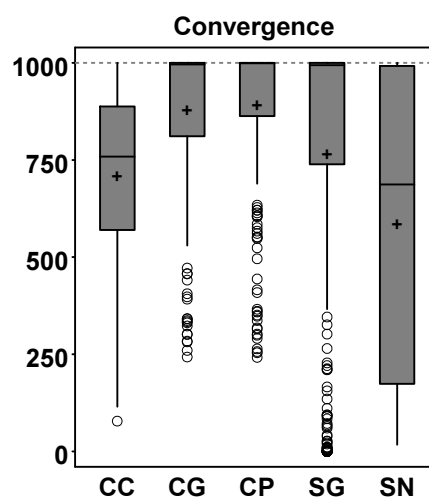


Figure 4. Distribution of number of converged simulation runs across the 251 simulation scenarios which gave complete results for all five procedures, given as boxplots. The following abbreviations are used: CC=Clayton copula, CG=Gauss copula, CP=Plackett copula, SG=Standard model estimated by penalized quasi-likelihood in SAS GLIMMIX, and SN=Standard model estimated by Gaussian quadrature in SAS NLMIXED.