Research Article

Statistics in Medicine

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-analyzing 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 and a larger flexibility for the association structure of sensitivity and specificity. In a simulation study, which compares three copula models and two implementations of the standard model, the Plackett and the Gauss copula do rarely perform worse, but frequently better than the standard model. 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-analyzing 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. 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)*

Statistics
in Medicine

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 not that uncommon.

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 $i = (1, \ldots, I)$) 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:

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

2.1. The standard model

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

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

with

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

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 estimate 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. Moreover, Chu et al. [11] generalized the model by allowing different link functions (probit and complementary log-log) in addition to the the standard logit function.

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 [0,1] scale. To be concrete, we assume them to be Beta distributed with parameters α and β,

$$Se \sim \text{Beta}(\alpha_{Se}, \beta_{Se}), \quad Sp \sim \text{Beta}(\alpha_{Sp}, \beta_{Sp}), \tag{4}$$

and density

$$f(u; \alpha_{Se}, \beta_{Se}) = \frac{\Gamma(\alpha_{Se} + \beta_{Se})}{\Gamma(\alpha_{Se})\Gamma(\beta_{Se})} u^{\alpha_{Se}-1}(1-u)^{\beta_{Se}-1} \tag{5}$$

for Se, and analogously defined for Sp. The expectations $(\alpha_*/(\alpha_* + \beta_*))$ of this two Beta distributions describe the meta-analytic parameters of interest (Se, Sp), the variances $(\alpha_*\beta_*/(\alpha_* + \beta_* + 1)(\alpha_* + \beta_*)^2)$ 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:

$$P(TP_i = k) = \binom{TP_i + FN_i}{k} \frac{B(\alpha_{Se} + k, TP_i + FN_i - k + \beta_{Se})}{B(\alpha_{Se}, \beta_{Se})}, \tag{6}$$

Prepared using simauth.cls

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 and also to the standard model in (2) and (3), the resulting marginal distribution has still closed form [12].

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 [13] 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 [14] proved that there exists a function C with

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

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 [13], 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 accomplished 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 using 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 which calls for differentiation with respect to the counting measure. In the work here, we nevertheless use only the continuous likelihoods.

To quantify the association between X_1 and X_2, it was shown [15] 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

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

and

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

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 rich 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 [16], the Gauss, and the Plackett copula. These were chosen because they allow the full range of associations and have been used in applied research before. It should be noted that the list of Archimedean copulas in table 4.1 from Nelsen [13] includes two other copulas (4.2.16 and 4.2.17) that allow the full range of associations. These were included in preliminary analyzes of ours but were later omitted to reduce complexity.

Clayton copula The Clayton copula [16] is defined the following way:

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

Its density is given by

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

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 $\tau = \theta_C/(\theta_C + 2)$.

Statistics
in Medicine

O. Kuss, A. Hoyer, A. Solms

Gauss copula The Gauss copula is defined by

$$C_G(u_1, u_2, \Gamma) = \Phi_2(\Phi^{-1}(u_1), \Phi^{-1}(v_2)|\Gamma) \tag{12}$$

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} refers to the distribution function of the standard (univariate) normal distribution. The density of the Gauss copula can be shown to be [17]

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

with $\mathbf{q} = (q_1, q_2)^{\mathrm{T}}$ 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_G (the off-diagonal element of Γ) is not straightforward. Song [17] 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 [18]. Later Nelsen [13] 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

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

and the density is

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

Sperman's ρ equals

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

Nelsen [13] 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. Aiming for a rather complete picture, we varied the following six factors:

- The true model (standard model, or one of the three copula introduced above, that is, Clayton, Gauss, or Plackett)
- The true sensitivity and specificity (70%/70%, 90%/70%, or 90%/90%)
- The true variance of sensitivity and specificity (small or large, where for the standard model these values were 0.15/0.75 on the logit scale, and for the copula models 0.01/0.05 on the [0,1] scale)
- The true association between sensitivity and specificity (none, small or large, where these categories stand for correlations of approximately 0, -0.2, or -0.8 on the respective scale)
- The number of studies in the respective meta-analysis (10 or 50)
- The number of observations in the respective study group of healthy and diseased probands (20 or 100)

Combination of the six design factors resulted in a total of 288 simulation scenarios. For each fixed scenario, 1000 meta-analyses were generated. True values for the different scenarios were chosen a priori and after analyzing a random sample of 15 meta-analyzes compiled by Menke [5]. To this task, the original data from these meta-analyses were downloaded and sensitivity, specificity, their variances and their association were computed. Values that we felt were typical from these meta-analyses were then fixed as true values for the simulation.

3.1. Data generation

To generate the actually observed data for a single study, a pair of random numbers was generated depending on the underlying true model (and, of course, the other fixed parameter for the respective scenario). In the case of the standard model assumed to be true the two random numbers were generated from a bivariate normal distribution with the VNORMAL call in SAS/IML. In the case of the copula models assumed to be true the two numbers (which were then bivariate uniform) were generated by using the standard algorithm for Archimedean copulas as given in Frees/Valdez [19] in case of the Clayton copula, by using the SOLVE statement from the SAS MODEL procedure in case of the Gauss copula, or by using the algorithm as proposed by Johnson [20] in case of the Plackett copula. A logit transformation (standard model) or an inverse Beta-transformation (copula models) of the respective bivariate pair then produced valid values for sensitivity and specificity. These numbers were multiplied with the number of observations in the respective group and then rounded to give observed numbers of true positives and true negatives. Groups of diseased and healthy people in the single studies were always assumed to be of equal size, mirroring the familiar case-control design in diagnostic accuracy studies.

3.2. Estimation

From each of the generated meta-analyses we estimated five parameters: sensitivity, specificity, their variances and their association. For each of the parameters we recorded the estimated value, its standard error, and a 95% confidence interval, where with respect to confidence intervals we made sure that limits were always within the allowed range (e.g. within [0,1] for sensitivity and specificity or [-1,1] for a correlation coefficient).

We compared five estimation methods, two assuming the standard model (Gaussian quadrature with SAS NLMIXED, SN, and Penalized quasi-likelihood with SAS GLIMMIX, SG) and one for each copula model (where we used maximum likelihood estimation with SAS NLMIXED). To enable a fair comparison between models, all SAS procedures were run with the default options. Starting values (which can and in general also should 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 correlations and transformed as appropriate for the respective model.

Following the current guideline of Burton et al. [21] we used bias, mean squared error (MSE) and empirical coverage (to the 95%-level) as outcomes to compare the five different estimation methods. To assess the numerical robustness of the different procedures, we also report the number of non-missing values for (1) the estimated parameter and (2) its standard error. The simulation programme was written in SAS.

We did make no allowance for the number of non-missing values when comparing models with respect to bias, MSE and coverage. This takes us on the 'safe side' when propagating copula models due to the following consideration: Under the plausible assumption that non-convergence occurs in more challenging situations where worse values of bias, MSE and coverage are expected, results for estimation methods with less convergence will be somewhat biased towards positive results. As we observe superior results for convergence in the copula models, the copula models are somewhat penalized with respect to bias, MSE and coverage.

3.3. Results

We restrict the reporting of results to sensitivity and specificity which were a priori defined as the parameters of main interest. We also omit all results for the large variance scenario because in this numerically more challenging situations the number of converged runs became too low and too diverse to enable a fair comparison between estimation methods. To additionally save space we omit all results from the Clayton copula (that is, all simulation scenarios where the true data were generated from the Clayton copula as well as all estimation results from this model). In terms of sample sizes of studies and probands we only report results for the situation with 10 studies and 100 probands per group in this paper as these mirror most closely our data example. Results for the other sample sizes scenarios are given in an online appendix.

PLACE TABLE 1 APPROXIMATELY HERE

In terms of bias, all methods perform similarly and have low bias when data were generated from the standard model. With data generated from the copula models, the estimation methods for the standard model (SG and SN) show some bias, especially in the situations with higher values for sensitivity and specificity, where bias is larger than 3%-points in some cases. Bias is largest when the data were generated from the Plackett copula. This can be partly explained by the unsymmetric nature of the Plackett copula where generated pairs of sensitivity and specificity show a rather erratic behavior, sometimes even leading to multimodal bivariate distributions of sensitivity and specificity. This unsymmetry is probably also the reason why biases for sensitivity and specificity from the same model diverge largely even in situations where they have the same true value of 70% or 90%. Both estimation methods assuming the standard model are not

Statistics
in Medicine

sensitive towards stronger associations between sensitivity and specificity, whereas with the copula estimation methods bias increases with increasing association, at least when data were generated from the copula models. The estimation method assuming the Plackett copula outperforms all other estimation methods, at least when data were generated form one of the copula models.

PLACE TABLE 2 APPROXIMATELY HERE

With regard to coverage, overall performance of the estimation methods is not entirely satisfactory as nearly all observed coverages are below the aspired 95%. Overall, methods perform best when the data were generated from the assumed model. Surprisingly low values (even down to below 10%) are found for the SG method, even with data generated from the standard model. Coverage worsens in most cases with higher values for true sensitivity and specificity and with stronger associations.

PLACE TABLE 3 APPROXIMATELY HERE

Considering mean squared error, all estimation methods perform rather similar when data were generated from the standard model or from the Gauss copula. The estimation methods assuming the standard model show some surprisingly large values when the Plackett copula with non-zero associations was used to generate the data.

PLACE TABLE 4 APPROXIMATELY HERE

In terms of numerical robustness the copula methods perform better than the methods for the standard model which is especially true for the most challenging situation of strong association. This behavior is expected as the estimation methods for the standard models estimate random effects models whereas the copula models essentially fit a bivariate distribution without random effects. Comparing Gaussian quadrature (SN) and PQL estimation (SG) we see that the latter is more robust which is expected as this method relies on fitting a linear mixed model for a pseudo-response as compared to a truly nonlinear model in the SN case. It should be noted that the standard model estimation methods do even have convergence problems when data were generated from the standard model.

4. Example

As an example we use the meta-analysis of Glas et al. [22] 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 5) was used at several instances in the methodical literature [7, 23], 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. [23] reported results for the standard model with SAS NLMIXED, however, only after trying a range of starting values. Glas et al. [22] 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 5 APPROXIMATELY HERE

In table 6 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 the estimated specificities which range from 81.9% from the Gauss copula to 91.2% from the standard model with Gaussian quadrature. The Plackett copula finds a specificity of 86.5%, which most closely approximates the original analysis. 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. The larger estimated values for specificity from the standard models mirror the results from the simulation study. In many situations (at least, when true data were generated from the copula models and especially from the asymmetric Plackett copula) we saw positive bias for the standard model estimation methods and thus an overestimation of parameter values.

PLACE TABLE 6 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 and a larger flexibility for the association structure of sensitivity and specificity. 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 and the Gauss copula do rarely perform worse, but frequently better than 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. [24], 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 [25] termed the corresponding class of distributions from the standard model 'logistic-normal distributions' and showed that these indeed yield slightly more flexible models. Still, both classes are restricted to 'U-shaped, J-shaped, and flattish to sharpish unimodal curves.' It should be noted, however, that the standard model in principle also allows for more complicated, like mixtures or fully nonparametric, distributions for the random effects which is not easily accomplished for by copula models. 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, 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.

It would certainly be interesting to have tools like model selection criterions or goodness-of-fit tests to compare the fits of the various models, between the copula and the standard model one one hand or between the different copula models on the other. As the likelihood functions between the standard model and the copulas models are completely different, we feel that there is no sensible definition of an AIC or a BIC or any other likelihood-based criterion, even the different copula models can not be compared by likelihood-based measures. The standard model does also suffer from this incomparableness, as it is not sensible to compare a Gaussian quadrature with a PQL fit. Possible solutions to the model selection problem are cross-validation or leave-one-out experiments to compare model fits. We also point to the vast literature of goodness-of-fit testing between copulas (see, e.g., [26]). However, these methods are mathematically complex and rarely used in applied work.

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 with our marginal Beta-binomial distributions which are discrete and should be differentiated with respect to the counting measure. Genest/Nešlehová [27] 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 that those for continuous marginals. We collected some limited information from simulating with the discrete likelihoods (data not shown) and found no relevant differences to the results from the continuous likelihoods. Instead and as expected due to the more complicated likelihoods, using the discrete likelihoods resulted in a somewhat 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. [24] 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 [28]. Only recently, we came across the problem of a meta-analytic comparison of two diagnostic tests to a common gold-standard that are applied within the same patient [29]. This problem might be modelled by a generalized linear mixed model with a quadrivariate response

Statistics
in Medicine

as well as by a quadrivariate copula. 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 for others 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*. 2012; 21(6):621-33.
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*. 2011 Dec 4. [Epub ahead of print]
11. Chu H, Guo H, Zhou Y. Bivariate Random Effects Meta-analysis of Diagnostic Studies Using Generalized Linear Mixed Models. *Medical Decision Making*. 2010;30:499–508.
12. Molenberghs G, Verbeke G. *Models for Discrete Longitudinal Data*. Springer: New York, 2006.
13. Nelsen RB. *An Introduction to Copulas*. Springer: New York, 2006.
14. 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.
15. Schweizer B, Wolff EF. On nonparametric measures of dependence for random variables. *Annals of Statistics*. 1981; 9:879-885.
16. 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.
17. Song PXK. Multivariate Dispersion Models Generated From Gaussian Copula. *Scandinavian Journal of Statistics*. 2000; 27(2):305-320.
18. Plackett RL. A Class of Bivariate Distributions. *Journal of the American Statistical Association*. 1965; 60(310):516-522.
19. Frees EW, Valdez E. Understanding Relationships Using Copulas. *North American Actuarial Journal*. 1998; 2(1):1-25.
20. Johnson ME. *Multivariate statistical simulation*. Wiley: New York, 1987.
21. Burton A, Altman DG, Royston P, Holder RL. The design of simulation studies in medical statistics. *Statistics in Medicine*. 2006; 25(24):4279-92.
22. 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.
23. 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.
24. 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.
25. Aitchison J, Begg CB. Statistical Diagnosis When Basic Cases are Not Classified with Certainty. *Biometrika*. 1976; 63(1), 1-12.
26. Berg D. Copula goodness-of-fit testing : an overview and power comparison. *The European Journal of Finance*. 2009; 15(7/8):675-701.
27. Genest C, Nešlehová J. A Primer on Copulas for Count Data. *Astin Bulletin*. 2007; 37(2): 475-515.
28. Kao SC, Govindaraju RS. Trivariate statistical analysis of extreme rainfall events via the Plackett family of copulas. *Water Resources Research*. 2008; 44:W02415.
29. Siadaty MS, Philbrick JT, Heim SW, Schectman JM. Repeated-measures modeling improved comparison of diagnostic tests in meta-analysis of dependent studies. *Journal of Clinical Epidemiology*. 2004; 57(7):698-711.

Statistics
in Medicine

O. Kuss, A. Hoyer, A. Solms

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;
     PARMS b_se=1.136 b_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(b_se)/(1 + EXP(b_se));
     pi_sp = EXP(b_sp)/(1 + EXP(b_sp));

   * Link the transformed binomial probabilities with the parameters from the Beta distributions;
     alpha_se =     pi_se *(1-rho_se)/rho_se ;
     beta_se  = (1- pi_se)*(1-rho_se)/rho_se ;
     alpha_sp =     pi_sp *(1-rho_sp)/rho_sp ;
     beta_sp  = (1- pi_sp)*(1-rho_sp)/rho_sp ;

   * Beta-binomial loglikelihood functions for sensitivity and specificity;
     ll_se =  LGAMMA(s+1)+LGAMMA(tp+alpha_se)+LGAMMA(s-tp+beta_se)+LGAMMA(alpha_se+beta_se)
             -LGAMMA(tp+1)-LGAMMA(s-tp+1)-LGAMMA(s+alpha_se+beta_se)-LGAMMA(alpha_se)-LGAMMA(beta_se);
     ll_sp = LGAMMA(h+1)+LGAMMA(tn+alpha_sp)+LGAMMA(h-tn+beta_sp)+LGAMMA(alpha_sp+beta_sp)
            -LGAMMA(tn+1)-LGAMMA(h-tn+1)-LGAMMA(h+alpha_sp+beta_sp) -LGAMMA(alpha_sp)-LGAMMA(beta_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_sepdf = EXP(LGAMMA(s+1) + LGAMMA(i+alpha_se) + LGAMMA(s-i+beta_se) + LGAMMA(alpha_se+beta_se)
                    - LGAMMA(i+1) - LGAMMA(s-i+1) - LGAMMA(s+alpha_se+beta_se)
                    - LGAMMA(alpha_se) - LGAMMA(beta_se));
        F_se = F_se + f_sepdf;
     end;
     F_sp = 0;
     do i=0 to tn;
        f_sppdf = EXP(LGAMMA(h+1) + LGAMMA(i+alpha_sp) + LGAMMA(h-i+beta_sp) + LGAMMA(alpha_sp+beta_sp)
                    - LGAMMA(i+1) - LGAMMA(h-i+1) - LGAMMA(h+alpha_sp+beta_sp)
                    - LGAMMA(alpha_sp) - LGAMMA(beta_sp));
        F_sp = F_sp + f_sppdf;
     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 log likelihood function;
     ll = ll_se + ll_sp + ll_plackett;

     MODEL tp ~ GENERAL(ll);

   * Compute z-transformed Plackett's rho;
     ESTIMATE "z-transformed Plackett's Rho" ARTANH((theta+1)/(theta-1)
                                          - ((2*theta)/(theta-1)**2)*LOG(theta));
```


Prepared using simauth.cls

```
      ODS OUTPUT ParameterEstimates=ParameterEstimatesCP(WHERE=(Parameter IN ("b_se","b_sp"))
                                  KEEP=Parameter Estimate Lower Upper)
                 AdditionalEstimates=AdditionalEstimatesCP(KEEP=Estimate Lower Upper);

  RUN;

  DATA TransformedParameterEstimatesCP(DROP=Parameter Estimate Lower Upper);
       SET ParameterEstimatesCP;
       IF Parameter="b_se" THEN NewParameter="Sensitivity";
       IF Parameter="b_sp" THEN NewParameter="Specificity";
       Value = EXP(Estimate)/(1 + EXP(Estimate));
       Lower_95CI = EXP(Lower)/(1 + EXP(Lower));
       Upper_95CI = EXP(Upper)/(1 + EXP(Upper));
       TITLE"Plackett-Copula, Estimates for Sensitivity and Specificity (with 95\%-CIs)";
  RUN;
  PROC PRINT DATA=TransformedParameterEstimatesCP;RUN;


* Retransform the z-transformed estimator for Plackett's rho;
  DATA RetransformedRhoCP(DROP=Estimate Lower Upper);
       SET AdditionalEstimatesCP;
       Placketts_Rho=(EXP(2*Estimate)-1)/(EXP(2*Estimate)+1);
       Lower_95CI_Placketts_Rho=(EXP(2*Lower)-1)/(EXP(2*Lower)+1);
       Upper_95CI_Placketts_Rho=(EXP(2*Upper)-1)/(EXP(2*Upper)+1);
       title"Plackett-Copula, Estimate for Plackett's rho (with 95\%-CIs)";
  RUN;
  PROC PRINT DATA=RetransformedRhoCP;RUN;
```

6.2. *R*

```
# Input telomerase data

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. Kavaler 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 the Plackett copula

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

  # Numbers of studies
  N  = dim(data)[1]

  se = as.numeric(parms["se"]); rho_se = as.numeric(parms["rho_se"])
  sp = as.numeric(parms["sp"]); rho_sp = as.numeric(parms["rho_sp"])
  theta = as.numeric(exp(parms["ltheta"]))

  # 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))
```

Statistics in Medicine

O. Kuss, A. Hoyer, A. Solms

```
  pi_se  = exp(se)/(1+exp(se))
  pi_sp  = exp(sp)/(1+exp(sp))

  # Link the transformed binomial probabilities with the parameters from the Beta distributions
  alpha_se =     pi_se *(1-rho_se)/rho_se
  beta_se  = (1- pi_se)*(1-rho_se)/rho_se
  alpha_sp =     pi_sp *(1-rho_sp)/rho_sp
  beta_sp  = (1- pi_sp)*(1-rho_sp)/rho_sp

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

  # Beta-binomial loglikelihood functions for sensitivity and specificity
  ll_se = lgamma(s+1)+lgamma(tp+alpha_se)+lgamma(s-tp+beta_se)+lgamma(alpha_se+beta_se)-
    lgamma(tp+1)-lgamma(s-tp+1)-lgamma(s+alpha_se+beta_se)-lgamma(alpha_se)-lgamma(beta_se)
  ll_sp = lgamma(h+1)+lgamma(tn+alpha_sp)+lgamma(h-tn+beta_sp)+lgamma(alpha_sp+beta_sp)-
    lgamma(tn+1)-lgamma(h-tn+1)-lgamma(h+alpha_sp+beta_sp)-lgamma(alpha_sp)-lgamma(beta_sp)

  # Define the distribution functions of the Beta-binomial distribution
  F_se = rep(NA, len = N); F_sp = rep(NA, len = N); ll_copula = rep(NA, len = N)
  for (j in 1:N)
  {
      f_se = rep(NA, len = tp[j])
      for (i in 1:tp[j]) f_se[i] = exp(lgamma(s[j]+1)+lgamma(i+alpha_se)+lgamma(s[j]-i+beta_se)+
                                       lgamma(alpha_se+beta_se)-lgamma(i+1)-lgamma(s[j]-i+1)-
                                       lgamma(s[j]+alpha_se+beta_se)-lgamma(alpha_se)-lgamma(beta_se))
        F_se[j]      = sum(f_se)
      f_sp = rep(NA, len = tn[j])
      for (i in 1:tn[j]) f_sp[i] = exp(lgamma(h[j]+1)+lgamma(i+alpha_sp)+lgamma(h[j]-i+beta_sp)+
                                       lgamma(alpha_sp+beta_sp)-lgamma(i+1)-lgamma(h[j]-i+1)-
                                       lgamma(h[j]+alpha_sp+beta_sp)-lgamma(alpha_sp)-lgamma(beta_sp))
      F_sp[j]      = sum(f_sp)

      # Define the Plackett density;
      if (theta != 1){
      ll_copula[j] = -3/2*log((1+(theta-1)*(F_se[j]+F_sp[j]))**2-4*theta*(theta-1)*F_se[j]*F_sp[j])+
        log(theta*(1+(theta-1)*(F_se[j]+F_sp[j]-2*F_se[j]*F_sp[j])))
      } else { ll_copula[j] = F_se[j]*F_sp[j] }

  }

  # Collect the elements of the log likelihood function
  logLL  = ll_se + ll_sp + ll_copula

  # Return -log likelihood
  -sum(logLL)
}

# Transform the estimation results

results = function(est, conf.level = 0.95, df)
  # 1.) est = output of optim fitting procedure (with hessian = TRUE)
  # 2.) conf.level, e.g. = 0.95 for 95\% Confidence Intervals
  # 3.) df = degree of freedom for students-t distribution
  #     (here numbers of studies)
{
  # Original parameter
  se        = as.numeric(est$par["se"])
  sp        = as.numeric(est$par["sp"])
  ltheta    = est$par["ltheta"]
  theta     = as.numeric(exp(ltheta))
  pi_se     = exp(se)/(1+exp(se))
  pi_sp     = exp(sp)/(1+exp(sp))
  # Spearman's  rho measurement of concordance
  rho = 0
  if (theta != 0) rho = (theta+1)/(theta-1)-2*theta/(theta-1)**2*log(theta)
  # Fisher's Z-transformation of rho
  rho_z     = atanh(rho)
```


Prepared using simauth.cls

```
  # Inverse Fisher information matrix
  invFisher = solve(est$hessian)
  # Standard errors of transformed-scale parameter
  SEs       = sqrt(diag(invFisher))
  names(SEs)= rownames(invFisher)

  # Assuming a normal distributed estimator (unknown mean and unknown variance)
  crit      = qt(1-(1-conf.level)/2, df = df-1)
  # Confidence levels of transformed-scale parameter
  CIs       = rbind(lb=c(se = se, sp = sp, est$par["ltheta"])-crit*SEs[c("se","sp","ltheta")],
                    ub=c(se = se, sp = sp, est$par["ltheta"])+crit*SEs[c("se","sp","ltheta")])

  # Standard errors of original-scale parameter via Delta-Method
  SEs["Sens"]  = pi_se*(1-pi_se)*SEs["se"]
  SEs["Spec"]  = pi_sp*(1-pi_sp)*SEs["sp"]
  SEs["theta"] = theta*SEs["ltheta"]
  grad1        = deriv(~(exp(ltheta)+1)/(exp(ltheta)-1)-2*exp(ltheta)/(exp(ltheta)-1)**2*ltheta,
                       "ltheta", function.arg = TRUE)
  fact1 = as.numeric(attr(grad1(ltheta = ltheta), "gradient"))
  SEs["rho"]   = fact1*SEs["ltheta"]
  SEs_rho_z    = 1/(1-rho**2)*SEs["rho"]

  # Confidence levels of original-scale parameter
  CI_ltheta    = as.numeric(est$par["ltheta"]+crit*c(-1,1)*SEs["ltheta"])
  grad2        = deriv(~.5*log((1+(exp(ltheta)+1)/(exp(ltheta)-1)-2*exp(ltheta)/(exp(ltheta)-1)**2*ltheta)/
                               (1-(exp(ltheta)+1)/(exp(ltheta)-1)+2*exp(ltheta)/(exp(ltheta)-1)**2*ltheta)),
                       "ltheta", function.arg = TRUE)
  fact2        = as.numeric(attr(grad2(ltheta = ltheta), "gradient"))

  CI_rho_z     = rho_z+crit*c(-1,1)*fact2*SEs["ltheta"]
  CI_rho       = tanh(CI_rho_z)
  CI_se        = exp(CIs[,"se"])/(1+exp(CIs[,"se"]))
  CI_sp        = exp(CIs[,"sp"])/(1+exp(CIs[,"sp"]))

  out = data.frame(rbind(
    par   = c(pi_se, pi_sp, rho),
    SE    = c(SEs["Sens"], SEs["Spec"], SEs["rho"]),
    CI_lb = c(CI_se[1], CI_sp[1], CI_rho[1]),
    CI_up = c(CI_se[2], CI_sp[2], CI_rho[2])))
    names(out) = c("Sens","Spec","rho")
  out
}

# Give the parameters to be estimated along with their starting 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 = optim(par     = initial,
              fn      = logLL.BivBetaBin,
              data    = telomerase,
              hessian = TRUE)
  res = results(est = fit, conf.level = 0.95, df=dim(telomerase)[1])
  print(round(res,4))
```

Statistics in Medicine

O. Kuss, A. Hoyer, A. Solms

Table 1. Bias (in percentage points) for sensitivity and specificity on the [0,1] scale for the various combinations of the underlying true model, true values for sensitivity and specificity, true associations between sensitivity and specificity, and the four estimation methods estimation methods. Results reported here are from the situation where the number of studies in the respective meta-analysis was set to 10 and had 100 observations in the respective study group of healthy and diseased probands. Abbreviations: CG=Gauss copula, CP=Plackett copula, SG=Penalized quasi-likelihood fit assuming the standard model, SN=Gaussian quadrature assuming the standard model

		Estimation method							
True sensitivity and specificity	True association between sensitivity and specificity	CG		CP		SG		SN	
		Se	Sp	Se	Sp	Se	Sp	Se	Sp
		True model: Gauss copula							
70%/70%	none	0.04	-0.02	0.03	-0.04	0.60	0.54	0.71	0.65
	weak	0.23	-0.11	0.22	-0.13	0.69	0.35	0.82	0.42
	strong	0.64	0.28	0.62	0.24	0.89	0.51	1.26	0.42
90%/70%	none	0.17	-0.02	0.16	-0.02	2.88	0.55	3.26	0.67
	weak	0.34	0.31	0.32	0.16	2.98	0.55	3.38	0.64
	strong	0.44	1.70	0.30	1.00	3.03	0.73	3.33	0.85
90%/90%	none	-0.08	-0.07	-0.09	-0.05	2.73	2.72	3.16	3.14
	weak	0.50	0.34	0.33	0.20	2.95	2.92	3.35	3.33
	strong	1.66	1.76	0.81	0.97	2.78	2.95	3.02	3.11
		True model: Plackett copula							
70%/70%	none	-0.13	0.01	-0.15	0.00	0.47	0.58	0.59	0.66
	weak	0.92	-0.15	0.14	-0.13	0.38	5.06	0.44	5.66
	strong	1.18	4.57	-0.37	3.66	0.68	7.80	0.41	10.12
90%/70%	none	0.05	-0.05	0.04	-0.05	2.84	0.54	3.25	0.66
	weak	1.07	0.56	0.18	-0.62	2.78	3.85	3.16	4.20
	strong	1.39	6.59	-0.54	4.94	2.64	8.58	2.89	10.12
90%/90%	none	0.05	0.03	0.06	-0.01	2.70	2.80	3.10	3.21
	weak	1.79	-5.68	1.08	-7.05	2.99	2.44	3.35	2.75
	strong	2.86	2.45	1.52	1.21	3.13	4.85	3.24	5.32
		True model: Standard model							
70%/70%	none	-0.61	-0.53	-0.63	-0.55	-0.32	-0.28	-0.28	-0.13
	weak	-0.56	-0.64	-0.54	-0.60	-0.35	-0.45	-0.31	-0.45
	strong	-0.14	-0.32	-0.09	-0.30	-0.14	-0.32	-0.10	-0.38
90%/70%	none	-0.66	-0.83	-0.64	-0.80	-0.46	-0.50	-0.51	-0.47
	weak	-0.54	-0.34	-0.54	-0.39	-0.43	-0.29	-0.40	-0.10
	strong	-0.47	0.40	-0.42	0.42	-0.43	0.16	-0.32	0.35
90%/90%	none	-0.69	-0.72	-0.69	-0.68	-0.45	-0.52	-0.52	-0.58
	weak	-0.56	-0.54	-0.54	-0.53	-0.45	-0.48	-0.45	-0.50
	strong	-0.26	-0.13	-0.18	-0.09	-0.52	-0.21	-0.06	0.12

Table 2. Empirical coverage (in %) for the 95% confidence intervals for sensitivity and specificity on the [0,1]scale for the various combinations of the underlying true model, true values for sensitivity and specificity, true associations between sensitivity and specificity, and the four estimation methods estimation methods. Results reported here are from the situation where the number of studies in the respective meta-analysis was set to 10 and had 100 observations in the respective study group of healthy and diseased probands. Abbreviations: CG=Gauss copula, CP=Plackett copula, SG=Penalized quasi-likelihood fit assuming the standard model, SN=Gaussian quadrature assuming the standard model

True sensitivity and specificity	True association between sensitivity and specificity	Estimation method							
		CG		CP		SG		SN	
		Se	Sp	Se	Sp	Se	Sp	Se	Sp
True model: Gauss copula									
70%/70%	none	93.3	93.5	94.0	93.6	92.7	92.9	94.0	94.3
	weak	93.4	94.1	92.8	94.0	93.2	93.4	94.3	95.2
	strong	92.8	92.8	91.7	93.2	89.3	89.9	90.3	89.3
90%/70%	none	92.1	93.5	92.0	94.1	82.1	94.0	83.7	95.0
	weak	93.7	92.5	93.1	93.3	80.7	92.1	83.6	93.5
	strong	92.0	89.3	92.0	92.5	77.7	87.7	88.3	92.4
90%/90%	none	93.8	92.5	93.8	93.4	85.2	84.0	87.5	86.3
	weak	92.6	92.0	92.7	92.2	82.9	84.1	85.6	85.8
	strong	88.3	86.4	92.0	90.4	73.5	70.8	86.5	84.4
True model: Plackett copula									
70%/70%	none	93.7	92.3	93.2	93.1	93.2	92.3	93.8	94.0
	weak	91.2	94.9	92.3	94.1	93.1	94.1	95.2	97.7
	strong	90.4	95.8	93.2	96.4	78.2	76.3	95.0	94.5
90%/70%	none	91.6	94.0	91.7	93.8	81.9	93.5	84.7	94.8
	weak	88.1	95.0	92.1	92.7	83.4	93.3	86.7	98.3
	strong	86.9	90.9	92.0	93.1	62.7	66.3	88.6	94.2
90%/90%	none	92.2	93.9	92.7	94.4	84.5	85.1	86.6	87.0
	weak	88.0	63.6	91.5	55.3	85.2	73.3	87.7	77.2
	strong	73.8	75.7	86.7	77.9	41.7	33.1	78.0	65.4
True model: Standard model									
70%/70%	none	93.4	92.8	93.4	92.9	94.7	93.7	95.7	95.5
	weak	94.3	93.3	94.3	92.1	94.4	92.4	94.7	94.6
	strong	93.5	94.2	92.4	90.8	86.7	87.3	69.8	82.7
90%/70%	none	79.5	91.4	78.6	91.7	92.6	90.0	95.3	91.7
	weak	80.2	92.3	81.1	92.6	88.7	88.0	93.8	90.3
	strong	81.1	89.9	81.1	90.3	35.3	34.6	77.8	67.6
90%/90%	none	81.6	77.0	77.7	75.8	88.1	85.8	95.1	91.2
	weak	80.7	80.2	77.3	79.7	83.3	83.9	92.4	91.8
	strong	85.8	85.0	84.3	83.0	8.7	9.0	84.3	74.4

Statistics in Medicine

O. Kuss, A. Hoyer, A. Solms

Table 3. Mean squared error (multiplied by 100) for the estimation of sensitivity and specificity on the [0,1] scale for the various combinations of the underlying true model, true values for sensitivity and specificity, true associations between sensitivity and specificity, and the four estimation methods estimation methods. Results reported here are from the situation where the number of studies in the respective meta-analysis was set to 10 and had 100 observations in the respective study group of healthy and diseased probands. Abbreviations: CG=Gauss copula, CP=Plackett copula, SG=Penalized quasi-likelihood fit assuming the standard model, SN=Gaussian quadrature assuming the standard model

True sensitivity and specificity	True association between sensitivity and specificity	Estimation method							
		CG		CP		SG		SN	
		Se	Sp	Se	Sp	Se	Sp	Se	Sp
True model: Gauss copula									
70%/70%	none	0.10	0.10	0.11	0.11	0.11	0.11	0.11	0.12
	weak	0.10	0.10	0.11	0.11	0.10	0.10	0.10	0.10
	strong	0.11	0.11	0.12	0.12	0.10	0.11	0.11	0.11
90%/70%	none	0.10	0.10	0.15	0.18	0.10	0.10	0.10	0.11
	weak	0.09	0.10	0.16	0.18	0.11	0.10	0.11	0.11
	strong	0.09	0.10	0.16	0.18	0.11	0.10	0.10	0.11
90%/90%	none	0.10	0.10	0.14	0.17	0.10	0.11	0.15	0.17
	weak	0.09	0.10	0.15	0.18	0.10	0.10	0.15	0.18
	strong	0.08	0.08	0.15	0.16	0.09	0.09	0.16	0.18
True model: Plackett copula									
70%/70%	none	0.11	0.11	0.11	0.12	0.10	0.10	0.11	0.11
	weak	0.11	0.11	0.11	0.11	1.05	1.07	1.53	1.85
	strong	0.11	0.11	0.11	0.11	0.57	0.46	1.13	1.60
90%/70%	none	0.11	0.11	0.15	0.18	0.11	0.11	0.11	0.11
	weak	0.10	0.10	0.15	0.18	1.02	1.16	1.66	2.02
	strong	0.09	0.12	0.15	0.16	0.70	0.58	1.25	1.56
90%/90%	none	0.10	0.10	0.14	0.16	0.10	0.10	0.14	0.17
	weak	0.10	0.09	0.16	0.18	1.61	2.01	0.46	0.65
	strong	0.13	0.08	0.17	0.20	0.24	0.32	0.31	0.40
True model: Standard model									
70%/70%	none	0.07	0.07	0.07	0.07	0.07	0.07	0.06	0.06
	weak	0.06	0.07	0.06	0.06	0.07	0.07	0.07	0.07
	strong	0.06	0.06	0.06	0.05	0.06	0.07	0.07	0.06
90%/70%	none	0.02	0.02	0.01	0.02	0.08	0.08	0.07	0.07
	weak	0.02	0.02	0.01	0.01	0.07	0.07	0.07	0.07
	strong	0.02	0.02	0.01	0.02	0.06	0.07	0.06	0.08
90%/90%	none	0.02	0.02	0.01	0.01	0.02	0.02	0.02	0.02
	weak	0.02	0.02	0.01	0.01	0.02	0.02	0.01	0.01
	strong	0.01	0.01	0.02	0.01	0.01	0.01	0.02	0.01

Table 4. Number of non-missing values (XXX/ YYY) for the estimated parameter (XXX) and the corresponding standard error (YYY) from 1000 simulation runs for the various combinations of the underlying true model, true values for sensitivity and specificity, true associations between sensitivity and specificity, and the four estimation methods estimation methods. Results reported here are from the situation where the number of studies in the respective meta-analysis was set to 10 and had 100 observations in the respective study group of healthy and diseased probands. Abbreviations: CG=Gauss copula, CP=Plackett copula, SG=Penalized quasi-likelihood fit assuming the standard model, SN=Gaussian quadrature assuming the standard model

True sensitivity and specificity	True association between sensitivity and specificity	Estimation method: CG Se	CG Sp	CP Se	CP Sp	SG Se	SG Sp	SN Se	SN Sp
		True model: Gauss copula							
70%/70%	none	999/ 995	999/ 993	991/ 987	991/ 989	1000/1000	1000/1000	955/949	955/950
	weak	995/ 993	995/ 991	992/ 989	992/ 991	999/ 999	999/ 999	930/926	930/925
	strong	997/ 994	997/ 990	991/ 989	991/ 987	968/ 968	968/ 968	326/292	326/291
90%/70%	none	997/ 997	997/ 994	998/ 998	998/ 995	1000/1000	1000/1000	967/966	967/967
	weak	991/ 990	991/ 985	994/ 992	994/ 987	998/ 998	998/ 998	954/952	954/950
	strong	980/ 979	980/ 968	996/ 995	996/ 993	935/ 935	935/ 935	530/478	530/485
90%/90%	none	999/ 999	999/ 999	999/ 999	999/ 999	994/ 994	994/ 994	990/989	990/987
	weak	1000/1000	1000/1000	1000/ 999	1000/1000	996/ 996	996/ 996	991/991	991/991
	strong	996/ 996	996/ 992	1000/1000	1000/1000	879/ 879	879/ 879	688/662	688/666
		True model: Plackett copula							
70%/70%	none	992/ 990	992/ 990	995/ 993	995/ 987	1000/1000	1000/1000	955/950	955/951
	weak	983/ 975	983/ 980	994/ 989	994/ 994	996/ 996	996/ 996	935/929	935/927
	strong	982/ 965	982/ 981	996/ 992	996/ 992	838/ 838	838/ 838	689/665	689/668
90%/70%	none	997/ 996	997/ 996	997/ 997	997/ 996	997/ 997	997/ 997	971/968	971/969
	weak	995/ 994	995/ 995	999/ 999	999/ 997	983/ 983	983/ 983	979/978	979/978
	strong	961/ 960	961/ 946	997/ 996	997/ 995	747/ 747	747/ 747	776/756	776/761
90%/90%	none	1000/1000	1000/1000	1000/1000	1000/1000	997/ 997	997/ 997	992/992	992/992
	weak	998/ 998	998/ 998	1000/1000	1000/1000	990/ 990	990/ 990	971/967	971/967
	strong	975/ 966	975/ 975	999/ 999	999/ 999	540/ 540	540/ 540	426/377	426/370
		True model: Standard model							
70%/70%	none	966/ 950	966/ 949	975/ 954	975/ 962	1000/1000	1000/1000	848/835	848/834
	weak	957/ 944	957/ 942	965/ 952	965/ 949	1000/1000	1000/1000	823/810	823/812
	strong	969/ 952	969/ 958	970/ 952	970/ 952	927/ 927	927/ 927	153/106	153/103
90%/70%	none	847/ 718	847/ 834	868/ 713	868/ 851	971/ 971	971/ 971	537/517	537/516
	weak	839/ 704	839/ 823	860/ 705	860/ 837	938/ 938	938/ 938	479/449	479/453
	strong	811/ 677	811/ 802	847/ 678	847/ 833	366/ 366	366/ 366	106/66	106/73
90%/90%	none	709/ 600	709/ 605	730/ 614	730/ 606	911/ 911	911/ 911	303/283	303/282
	weak	731/ 632	731/ 608	758/ 627	758/ 625	882/ 882	882/ 882	259/232	259/238
	strong	748/ 667	748/ 670	773/ 657	773/ 678	94/ 94	94/ 94	80/51	80/43

Statistics
in Medicine

O. Kuss, A. Hoyer, A. Solms

Table 5. Example data set for the telomerase data from Glas et al. [22]

Study	TP	FP	FN	TN
Ito 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 6. Results for the different estimation methods for the telomerase data from Glas et al. [22].

Method	Sensitivity (95%-CI)	Specificity (95%-CI)	Association (95%-CI)
Gauss	79.4% (74.8%, 83.4%)	81.9% (65.7%, 91.4%)	γ_G=-0.71 (-0.88, -0.38)
Plackett	77.6% (72.2%, 82.2%)	86.5% (70.0%, 94.6%)	ρ_S= -0.85 (-0.96, -0.56)
Standard model (GQ, NLMIXED)	76.7% (68.6%, 83.2%)	91.2% (70.8%, 97.8%)	ρ=-1.00 (–, –)
Standard model (PQL, GLIMMIX)	76.7% (68.8%, 83.1%)	90.4% (71.5%, 97.2%)	ρ=-1.00 (–, -0.33)
Standard model (MCMC) [7]	76.7% (69.8%, 82.6%)	90.9% (73.7%, 97.8%)	ρ=-0.88 (-0.99, -0.18)
Original analysis [22]	75% (71%, 79%)	86% (71%, 94%)	ρ'=-0.73 (–, –)

Prepared using simauth.cls

Statistics in Medicine

Additional simulation results from Kuss/Hoyer/Solms: Meta-analysis for diagnostic accuracy studies: A new statistical model using beta-binomial distributions and bivariate copulas

1. Number of studies: 10, Number of observations in the respective study group: 20

Table 1. Bias (in percentage points) for sensitivity and specificity on the [0,1]-scale for the various combinations of the underlying true model, true values for sensitivity and specificity, true associations between sensitivity and specificity, and the four estimation methods estimation methods. Results reported here are from the situation where the number of studies in the respective meta-analysis was set to 10 and had 20 observations in the respective study group of healthy and diseased probands. Abbreviations: CG=Gauss copula, CP=Plackett copula, SG=Penalized quasi-likelihood fit assuming the standard model, SN=Gaussian quadrature assuming the standard model

True sensitivity and specificity	True association between sensitivity and specificity	Estimated model: CG Se	CG Sp	CP Se	CP Sp	SG Se	SG Sp	SN Se	SN Sp
		True model: Gauss copula							
70%/70%	none	-0.09	-0.45	-0.35	-0.37	0.23	-0.03	0.14	-0.07
	weak	0.23	0.40	0.43	0.30	0.19	0.17	-0.01	-0.09
	strong	1.69	2.17	1.74	1.94	0.55	0.60	0.79	1.18
90%/70%	none	-0.41	-0.39	-0.26	-0.26	1.20	0.19	3.31	0.01
	weak	-0.27	0.42	0.01	0.43	1.28	0.11	3.80	-0.22
	strong	0.61	3.35	0.61	3.24	2.23	0.15	7.53	0.28
90%/90%	none	-0.97	-1.01	-0.07	-0.11	1.36	1.27	2.15	1.88
	weak	-0.46	-0.45	0.28	0.23	1.10	1.25	1.53	2.00
	strong	3.00	2.64	2.53	2.44	2.36	1.15	3.94	3.72
		True model: Plackett copula							
70%/70%	none	-0.32	-0.30	-0.20	-0.15	0.12	0.18	0.03	0.12
	weak	1.18	-0.07	1.00	0.54	0.31	2.68	-0.56	5.12
	strong	2.30	4.56	1.58	4.58	0.58	7.29	-1.15	13.95
90%/70%	none	-0.16	-0.48	-0.12	-0.21	1.33	0.21	3.67	-0.20
	weak	0.96	3.55	0.73	2.14	1.52	2.42	2.31	2.23
	strong	1.23	7.88	0.40	7.61	1.91	9.45	2.98	8.28
90%/90%	none	-1.06	-1.15	-0.09	-0.20	1.34	1.30	1.65	1.78
	weak	1.63	-4.08	1.87	-4.09	2.01	1.16	3.33	1.60
	strong	3.18	4.04	3.16	3.84	3.47	4.24	4.30	6.55
		True model: Standard model							
70%/70%	none	-0.91	-0.96	-0.75	-0.98	-0.52	-0.44	-0.15	-0.16
	weak	-0.19	-0.19	-0.21	-0.14	-0.32	-0.35	0.49	-0.62
	strong	1.27	1.41	1.26	1.31	-0.90	0.65	0.07	-0.68
90%/70%	none	-0.82	-0.99	-1.01	-1.00	-0.37	-0.74	-0.33	-0.19
	weak	-0.28	0.06	-0.71	-0.77	-0.44	-0.67	-0.38	-0.50
	strong	0.61	2.30	0.50	1.68	0.62	-1.56	0.29	-1.05
90%/90%	none	-0.90	-0.80	-1.99	-1.85	-0.76	-0.42	-0.37	-0.84
	weak	-0.37	-0.56	-1.37	-1.58	-0.39	-0.40	-0.16	-0.57
	strong	1.19	0.95	0.15	0.26	0.04	-0.29	-0.33	0.31

Statistics
in Medicine

Table 2. Empirical coverage (in %) for the 95% confidence intervals for sensitivity and specificity on the [0,1] scale for the various combinations of the underlying true model, true values for sensitivity and specificity, true associations between sensitivity and specificity, and the four estimation methods estimation methods. Results reported here are from the situation where the number of studies in the respective meta-analysis was set to 10 and had 20 observations in the respective study group of healthy and diseased probands. Abbreviations: CG=Gauss copula, CP=Plackett copula, SG=Penalized quasi-likelihood fit assuming the standard model, SN=Gaussian quadrature assuming the standard model

		Estimated model							
True sensitivity and specificity	True association between sensitivity and specificity	CG		CP		SG		SN	
		Se	Sp	Se	Sp	Se	Sp	Se	Sp
		True model: Gauss copula							
70%/70%	none	63.5	66.0	58.4	60.2	89.6	89.2	92.6	95.7
	weak	65.2	66.4	65.3	64.6	88.6	90.0	100.0	97.7
	strong	67.5	66.0	60.7	57.3	21.3	21.5	90.9	78.3
90%/70%	none	90.4	54.0	92.0	55.8	85.5	87.8	87.9	93.5
	weak	92.9	56.0	92.0	58.5	82.0	85.9	82.9	93.0
	strong	91.0	27.5	87.2	44.0	14.5	15.5	48.6	87.5
90%/90%	none	90.1	91.1	90.4	90.8	79.3	80.5	97.7	98.1
	weak	90.5	90.1	89.6	89.8	74.5	74.6	98.3	98.6
	strong	78.4	81.4	86.1	84.8	6.6	7.2	86.8	82.9
		True model: Plackett copula							
70%/70%	none	62.9	67.6	65.0	61.0	90.2	90.1	97.7	96.6
	weak	53.2	90.4	61.0	85.4	79.5	76.1	87.1	86.5
	strong	39.6	83.5	57.6	62.3	17.1	15.1	90.9	69.4
90%/70%	none	93.2	62.4	90.8	67.4	84.9	87.1	86.8	96.8
	weak	88.5	82.6	89.7	82.9	64.6	63.1	96.6	93.2
	strong	82.6	47.6	87.3	45.0	6.5	5.1	89.8	79.6
90%/90%	none	91.3	90.7	91.1	91.9	81.8	82.5	99.3	98.9
	weak	87.5	82.4	88.5	77.7	47.3	45.2	96.9	96.4
	strong	70.4	59.6	78.0	63.7	0.4	0.3	89.3	76.9
		True model: Standard model							
70%/70%	none	29.9	29.1	26.2	28.2	82.6	82.4	87.9	92.1
	weak	40.2	35.1	33.8	29.4	78.0	77.0	95.7	77.4
	strong	50.7	48.5	38.2	47.4	7.1	7.0	88.9	77.8
90%/70%	none	9.5	42.3	11.9	26.5	40.9	40.1	87.1	82.8
	weak	14.8	46.6	15.3	39.8	36.3	36.2	84.1	85.1
	strong	17.8	55.2	24.2	41.1	0.4	0.4	90.9	85.3
90%/90%	none	6.9	10.7	6.9	7.8	14.5	14.6	94.6	92.3
	weak	10.5	17.6	6.0	10.2	8.7	8.8	88.0	91.7
	strong	34.2	36.5	18.1	15.7	1.2	1.2	87.9	84.8

Statistics in Medicine

Table 3. Mean squared error (multiplied by 100) for the estimation of sensitivity and specificity on the [0,1] scale for the various combinations of the underlying true model, true values for sensitivity and specificity, true associations between sensitivity and specificity, and the four estimation methods estimation methods. Results reported here are from the situation where the number of studies in the respective meta-analysis was set to 10 and had 20 observations in the respective study group of healthy and diseased probands. Abbreviations: CG=Gauss copula, CP=Plackett copula, SG=Penalized quasi-likelihood fit assuming the standard model, SN=Gaussian quadrature assuming the standard model

True sensitivity and specificity	True association between sensitivity and specificity	Estimated model: CG Se	CG Sp	CP Se	CP Sp	SG Se	SG Sp	SN Se	SN Sp
True model: Gauss copula									
70%/70%	none	0.11	0.12	0.10	0.12	0.11	0.11	0.11	0.11
	weak	0.11	0.11	0.11	0.09	0.10	0.10	0.09	0.09
	strong	0.13	0.14	0.10	0.10	0.13	0.14	0.09	0.12
90%/70%	none	0.11	0.10	0.10	0.27	0.13	0.12	0.10	0.11
	weak	0.11	0.11	0.11	0.31	0.13	0.13	0.11	0.12
	strong	0.08	0.09	0.12	0.63	0.21	0.19	0.10	0.11
90%/90%	none	0.15	0.11	0.10	0.13	0.15	0.11	0.10	0.14
	weak	0.13	0.11	0.10	0.11	0.14	0.11	0.10	0.13
	strong	0.16	0.12	0.13	0.37	0.15	0.12	0.12	0.35
True model: Plackett copula									
70%/70%	none	0.12	0.12	0.11	0.11	0.12	0.11	0.11	0.10
	weak	0.13	0.13	0.11	0.15	0.72	0.66	0.75	2.60
	strong	0.15	0.13	0.13	0.17	0.44	0.40	0.88	2.80
90%/70%	none	0.10	0.11	0.10	0.29	0.12	0.11	0.10	0.11
	weak	0.11	0.10	0.12	0.18	0.69	0.74	0.79	1.55
	strong	0.09	0.09	0.12	0.29	0.88	0.77	1.18	1.61
90%/90%	none	0.15	0.11	0.10	0.11	0.15	0.11	0.10	0.12
	weak	0.13	0.11	0.13	0.21	1.15	1.01	0.43	0.81
	strong	0.25	0.14	0.13	0.32	0.32	0.26	0.23	0.61
True model: Standard model									
70%/70%	none	0.10	0.09	0.07	0.08	0.09	0.10	0.07	0.08
	weak	0.07	0.08	0.06	0.08	0.08	0.08	0.08	0.10
	strong	0.07	0.08	0.07	0.05	0.08	0.08	0.08	0.06
90%/70%	none	0.03	0.04	0.02	0.01	0.14	0.12	0.08	0.05
	weak	0.03	0.03	0.02	0.02	0.10	0.12	0.07	0.07
	strong	0.02	0.02	0.01	0.02	0.11	0.11	0.05	0.13
90%/90%	none	0.05	0.07	0.03	0.02	0.05	0.06	0.02	0.02
	weak	0.04	0.06	0.02	0.02	0.04	0.06	0.01	0.02
	strong	0.03	0.04	0.01	0.02	0.08	0.04	0.01	0.02

Statistics
in Medicine

Table 4. Number of non-missing values (XXX/ YYY) for the estimated parameter (XXX) and the corresponding standard error (YYY) from 1000 simulation runs for the various combinations of the underlying true model, true values for sensitivity and specificity, true associations between sensitivity and specificity, and the four estimation methods estimation methods. Results reported here are from the situation where the number of studies in the respective meta-analysis was set to 10 and had 20 observations in the respective study group of healthy and diseased probands. Abbreviations: CG=Gauss copula, CP=Plackett copula, SG=Penalized quasi-likelihood fit assuming the standard model, SN=Gaussian quadrature assuming the standard model

True sensitivity and specificity	True association between sensitivity and specificity	Estimated model: CG Se	CG Sp	CP Se	CP Sp	SG Se	SG Sp	SN Se	SN Sp
		True model: Gauss copula							
70%/70%	none	560/ 388	560/ 387	606/ 391	606/ 400	921/ 921	921/ 921	119/102	119/105
	weak	542/ 408	542/ 384	625/ 429	625/ 427	913/ 913	913/ 913	113/90	113/90
	strong	611/ 519	611/ 510	615/ 477	615/ 478	220/ 220	220/ 220	42/22	42/23
90%/70%	none	657/ 632	657/ 409	724/ 689	724/ 484	895/ 895	895/ 895	186/154	186/156
	weak	667/ 638	667/ 424	721/ 692	721/ 481	881/ 881	881/ 881	179/139	179/146
	strong	565/ 541	565/ 280	659/ 630	659/ 399	156/ 156	156/ 156	76/26	76/38
90%/90%	none	851/ 800	851/ 789	885/835	885/ 834	856/ 856	856/ 856	499/480	499/483
	weak	850/ 804	850/ 789	910/ 850	910/ 854	784/ 784	784/ 784	426/415	426/411
	strong	582/ 524	582/ 517	877/ 823	877/ 826	78/ 78	78/ 78	68/36	68/31
		True model: Plackett copula							
70%/70%	none	530/ 370	530/ 373	602/ 412	602/ 403	929/ 929	929/ 929	124/100	124/99
	weak	681/ 422	681/ 651	767/ 508	767/ 729	803/ 803	803/ 803	192/133	192/134
	strong	605/ 355	605/ 570	719/ 477	719/ 656	84/ 84	84/84	93/49	93/56
90%/70%	none	642/ 610	642/ 424	742/ 705	742/ 480	893/ 893	893/ 893	190/154	190/155
	weak	807/ 741	807/ 764	920/ 871	920/ 868	685/ 685	685/ 685	416/385	416/381
	strong	590/ 509	590/ 463	832/ 786	832/ 701	68/ 68	68/68	153/107	153/112
90%/90%	none	867/ 812	867/ 818	885/ 824	885/ 834	857/857	857/857	461/447	461/446
	weak	842/ 766	842/ 803	917/ 874	917/ 884	513/ 513	513/ 513	224/187	224/186
	strong	0/ 0	0/ 0	853/ 803	853/ 816	4/ 4	4/ 4	55/25	55/21
		True model: Standard model							
70%/70%	none	359/ 216	359/ 227	425/ 260	425/ 278	832/ 832	832/ 832	89/65	89/63
	weak	396/ 237	396/ 208	446/ 230	446/ 256	778/ 778	778/778	66/52	66/52
	strong	446/ 339	446/ 325	446/ 291	446/ 301	71/ 71	71/71	43/27	43/27
90%/70%	none	280/ 137	280/ 186	356/ 206	356/ 217	411/ 411	411/ 411	109/92	109/93
	weak	336/ 160	336/ 208	399/ 192	399/ 216	346/ 346	346/ 346	95/75	95/71
	strong	420/ 239	420/ 365	376/ 213	376/ 249	4/ 4	4/ 4	40/33	40/34
90%/90%	none	206/ 110	206/ 113	271/ 169	271/ 161	147/ 147	147/ 147	47/37	47/38
	weak	233/ 137	233/ 124	281/ 157	281/ 153	89/ 89	89/ 89	36/24	36/23
	strong	339/ 194	339/ 205	320/ 157	320/ 159	4/ 4	4/ 4	29/24	29/23

Prepared using simauth.cls

Statistics in Medicine

2. Number of studies: 50, Number of observations in the respective study group: 100

Table 5. Bias (in percentage points) for sensitivity and specificity on the [0,1] scale for the various combinations of the underlying true model, true values for sensitivity and specificity, true associations between sensitivity and specificity, and the four estimation methods estimation methods. Results reported here are from the situation where the number of studies in the respective meta-analysis was set to 50 and had 100 observations in the respective study group of healthy and diseased probands. Abbreviations: CG=Gauss copula, CP=Plackett copula, SG=Penalized quasi-likelihood fit assuming the standard model, SN=Gaussian quadrature assuming the standard model

		Estimated model							
True sensitivity and specificity	True association between sensitivity and specificity	CG		CP		SG		SN	
		Se	Sp	Se	Sp	Se	Sp	Se	Sp
		True model: Gauss copula							
70%/70%	none	0.02	-0.01	0.01	-0.02	0.60	0.60	0.77	0.76
	weak	0.07	0.09	0.06	0.08	0.57	0.60	0.74	0.77
	strong	0.36	0.44	0.35	0.43	0.64	0.72	0.72	0.86
90%/70%	none	0.00	-0.13	0.00	-0.10	3.01	0.49	3.55	0.66
	weak	-0.01	0.32	0.01	0.15	2.95	0.58	3.48	0.74
	strong	0.47	1.67	0.31	0.81	3.24	0.61	3.64	0.64
90%/90%	none	-0.02	0.01	0.02	0.04	3.03	3.03	3.57	3.57
	weak	0.33	0.25	0.22	0.11	3.05	2.99	3.58	3.52
	strong	1.81	1.79	0.86	0.83	3.18	3.15	3.50	3.46
		True model: Plackett copula							
70%/70%	none	-0.04	-0.00	-0.04	-0.00	0.56	0.62	0.73	0.79
	weak	1.09	-1.89	0.33	-1.93	0.54	5.56	0.69	6.46
	strong	1.18	4.40	-0.77	2.97	0.63	7.38	0.71	8.06
90%/70%	none	-0.00	-0.06	0.00	-0.06	2.99	0.55	3.53	0.72
	weak	1.21	-0.07	0.30	-1.41	3.12	5.70	3.60	6.59
	strong	1.30	6.25	-1.31	4.08	3.08	7.59	3.48	8.06
90%/90%	none	-0.01	-0.11	0.04	-0.07	3.02	2.99	3.56	3.53
	weak	1.71	-7.63	1.11	-8.94	3.11	3.90	3.57	4.80
	strong	2.91	2.10	1.47	0.88	3.36	5.06	3.57	5.53
		True model: Standard model							
70%/70%	none	-0.58	-0.57	-0.58	-0.57	-0.32	-0.31	-0.20	-0.18
	weak	-0.47	-0.44	-0.46	-0.44	-0.30	-0.27	-0.18	-0.15
	strong	-0.22	-0.19	-0.17	-0.14	-0.22	-0.19	0.05	-0.11
90%/70%	none	-0.52	-0.58	-0.52	-0.60	-0.45	-0.31	-0.38	-0.21
	weak	-0.51	-0.43	-0.51	-0.40	-0.47	-0.37	-0.40	-0.25
	strong	-0.37	0.26	-0.30	0.31	-0.23	-0.08	0.02	-0.06
90%/90%	none	-0.56	-0.58	-0.57	-0.59	-0.48	-0.49	-0.39	-0.40
	weak	-0.39	-0.46	-0.38	-0.45	-0.39	-0.46	-0.33	-0.40
	strong	-0.17	-0.14	-0.09	-0.07	-0.15	-0.42	0.06	-0.10

Statistics
in Medicine

Table 6. Empirical coverage (in %) for the 95% confidence intervals for sensitivity and specificity on the [0,1] scale for the various combinations of the underlying true model, true values for sensitivity and specificity, true associations between sensitivity and specificity, and the four estimation methods estimation methods. Results reported here are from the situation where the number of studies in the respective meta-analysis was set to 50 and had 100 observations in the respective study group of healthy and diseased probands. Abbreviations: CG=Gauss copula, CP=Plackett copula, SG=Penalized quasi-likelihood fit assuming the standard model, SN=Gaussian quadrature assuming the standard model

True sensitivity and specificity	True association between sensitivity and specificity	Estimated model: CG Se	CG Sp	CP Se	CP Sp	SG Se	SG Sp	SN Se	SN Sp
		True model: Gauss copula							
70%/70%	none	94.3	95.2	93.9	94.9	92.3	93.7	92.1	92.8
	weak	94.8	94.8	94.6	94.8	93.2	92.8	92.6	91.8
	strong	93.1	93.5	91.6	91.9	91.0	91.0	88.2	84.5
90%/70%	none	95.1	95.2	95.4	95.1	33.7	94.4	24.2	93.4
	weak	94.4	94.1	94.1	94.5	36.8	93.3	26.3	92.7
	strong	89.1	67.9	91.9	90.1	28.6	92.8	23.7	92.5
90%/90%	none	94.3	94.1	94.2	94.6	33.6	32.8	23.5	23.6
	weak	94.1	94.5	94.5	94.7	35.1	34.6	22.4	25.1
	strong	61.3	59.5	87.5	89.7	29.9	29.7	24.5	24.2
		True model: Plackett copula							
70%/70%	none	92.8	93.6	93.1	93.8	90.8	91.7	90.5	90.7
	weak	84.0	81.4	93.1	80.5	93.4	76.1	92.9	74.9
	strong	81.6	54.5	93.1	74.3	91.4	29.7	91.7	27.6
90%/70%	none	94.9	95.8	94.7	95.2	36.0	93.8	24.3	92.4
	weak	81.3	87.3	92.0	84.9	32.9	78.6	23.3	78.1
	strong	78.5	21.1	89.8	44.6	30.9	26.7	24.7	24.7
90%/90%	none	94.3	95.4	95.0	95.1	33.3	34.0	22.9	23.3
	weak	72.0	25.9	84.2	22.1	33.6	55.9	22.7	45.5
	strong	17.2	33.2	72.1	60.1	22.3	5.9	22.1	4.7
		True model: Standard model							
70%/70%	none	92.1	91.1	92.2	91.4	93.7	93.0	94.3	93.9
	weak	93.9	94.6	94.1	94.3	94.9	95.1	95.5	95.8
	strong	94.2	94.2	92.2	93.1	93.6	94.2	61.4	75.4
90%/70%	none	79.6	92.4	79.0	91.7	86.7	93.6	90.4	93.0
	weak	82.4	93.7	82.7	93.6	86.5	93.4	90.1	93.0
	strong	86.1	94.0	85.7	91.8	28.8	28.7	76.8	76.8
90%/90%	none	79.3	78.5	78.6	77.6	85.7	85.0	90.0	90.5
	weak	86.7	83.5	86.7	84.5	89.6	86.2	90.8	90.2
	strong	94.2	94.9	90.9	90.9	0.1	0.1	86.0	84.1

Statistics in Medicine

Table 7. Mean squared error (multiplied by 100) for the estimation of sensitivity and specificity on the [0,1] scale for the various combinations of the underlying true model, true values for sensitivity and specificity, true associations between sensitivity and specificity, and the four estimation methods estimation methods. Results reported here are from the situation where the number of studies in the respective meta-analysis was set to 50 and had 100 observations in the respective study group of healthy and diseased probands. Abbreviations: CG=Gauss copula, CP=Plackett copula, SG=Penalized quasi-likelihood fit assuming the standard model, SN=Gaussian quadrature assuming the standard model

		Estimated model							
True sensitivity and specificity	True association between sensitivity and specificity	CG		CP		SG		SN	
		Se	Sp	Se	Sp	Se	Sp	Se	Sp
True model: Gauss copula									
70%/70%	none	0.02	0.02	0.02	0.03	0.02	0.02	0.02	0.03
	weak	0.02	0.02	0.02	0.03	0.02	0.02	0.02	0.03
	strong	0.02	0.02	0.03	0.03	0.02	0.02	0.03	0.03
90%/70%	none	0.02	0.02	0.10	0.14	0.02	0.02	0.02	0.02
	weak	0.02	0.02	0.10	0.14	0.02	0.02	0.02	0.03
	strong	0.02	0.02	0.12	0.15	0.04	0.02	0.02	0.03
90%/90%	none	0.02	0.02	0.10	0.14	0.02	0.02	0.10	0.14
	weak	0.02	0.02	0.10	0.14	0.02	0.02	0.10	0.14
	strong	0.04	0.02	0.12	0.14	0.04	0.02	0.11	0.13
True model: Plackett copula									
70%/70%	none	0.02	0.02	0.03	0.03	0.02	0.02	0.03	0.03
	weak	0.03	0.02	0.02	0.03	0.31	0.31	0.57	0.73
	strong	0.03	0.03	0.03	0.03	0.27	0.14	0.65	0.77
90%/70%	none	0.02	0.02	0.10	0.14	0.02	0.02	0.02	0.02
	weak	0.03	0.02	0.11	0.14	0.21	0.27	0.58	0.74
	strong	0.03	0.04	0.11	0.13	0.45	0.23	0.69	0.78
90%/90%	none	0.02	0.02	0.10	0.14	0.02	0.02	0.10	0.14
	weak	0.04	0.03	0.11	0.14	0.91	1.18	0.19	0.26
	strong	0.09	0.03	0.13	0.14	0.09	0.07	0.27	0.32
True model: Standard model									
70%/70%	none	0.02	0.02	0.01	0.01	0.02	0.02	0.01	0.01
	weak	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	strong	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01
90%/70%	none	0.01	0.01	0.00	0.00	0.02	0.02	0.01	0.01
	weak	0.01	0.01	0.00	0.00	0.01	0.01	0.01	0.01
	strong	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01
90%/90%	none	0.01	0.01	0.00	0.00	0.01	0.01	0.00	0.00
	weak	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	strong	0.00	0.00	.	0.00	0.00	0.00	.	0.00

Statistics
in Medicine

Table 8. Number of non-missing values (XXX/ YYY) for the estimated parameter (XXX) and the corresponding standard error (YYY) from 1000 simulation runs for the various combinations of the underlying true model, true values for sensitivity and specificity, true associations between sensitivity and specificity, and the four estimation methods estimation methods. Results reported here are from the situation where the number of studies in the respective meta-analysis was set to 50 and had 100 observations in the respective study group of healthy and diseased probands. Abbreviations: CG=Gauss copula, CP=Plackett copula, SG=Penalized quasi-likelihood fit assuming the standard model, SN=Gaussian quadrature assuming the standard model

True sensitivity and specificity	True association between sensitivity and specificity	Estimated model: CG Se	CG Sp	CP Se	CP Sp	SG Se	SG Sp	SN Se	SN Sp
		True model: Gauss copula							
70%/70%	none	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000
	weak	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000
	strong	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	437/390	437/387
90%/70%	none	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000
	weak	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000
	strong	1000/1000	1000/ 999	1000/1000	1000/1000	987/ 987	987/ 987	779/ 757	779/757
90%/90%	none	1000/1000	1000/1000	999/ 999	999/ 999	1000/1000	1000/1000	1000/1000	1000/1000
	weak	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000
	strong	1000/ 999	1000/1000	1000/1000	1000/1000	988/ 988	988/ 988	985/984	985/985
		True model: Plackett copula							
70%/70%	none	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000
	weak	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000
	strong	1000/1000	1000/1000	1000/1000	1000/1000	996/ 996	996/996	968/967	968/965
90%/70%	none	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000
	weak	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000
	strong	992/ 992	992/ 980	1000/1000	1000/1000	984/ 984	984/984	998/998	998/998
90%/90%	none	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000
	weak	997/ 997	997/ 986	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000
	strong	997/ 993	997/ 933	1000/1000	1000/1000	936/ 936	936/ 936	610/564	610/563
		True model: Standard model							
70%/70%	none	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000
	weak	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000
	strong	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	1000/1000	107/70	107/61
90%/70%	none	966/931	966/966	969/936	969/969	1000/1000	1000/1000	913/900	913/898
	weak	975/927	975/974	973/925	973/973	999/999	999/999	851/841	851/833
	strong	978/959	978/978	961/929	961/959	302/302	302/302	98/56	98/56
90%/90%	none	945/919	945/923	962/944	962/935	1000/1000	1000/1000	778/772	778/767
	weak	935/901	935/912	942/899	942/914	1000/1000	1000/1000	570/555	570/553
	strong	957/949	957/936	938/903	938/897	1/1	1/1	76/41	76/44

Prepared using simauth.cls

Statistics in Medicine

3. Number of studies: 50, Number of observations in the respective study group: 20

Table 9. Bias (in percentage points) for sensitivity and specificity on the [0,1] scale for the various combinations of the underlying true model, true values for sensitivity and specificity, true associations between sensitivity and specificity, and the four estimation methods estimation methods. Results reported here are from the situation where the number of studies in the respective meta-analysis was set to 50 and had 20 observations in the respective study group of healthy and diseased probands. Abbreviations: CG=Gauss copula, CP=Plackett copula, SG=Penalized quasi-likelihood fit assuming the standard model, SN=Gaussian quadrature assuming the standard model

		Estimated model							
True sensitivity and specificity	True association between sensitivity and specificity	CG		CP		SG		SN	
		Se	Sp	Se	Sp	Se	Sp	Se	Sp
		True model: Gauss copula							
70%/70%	none	-0.11	-0.10	-0.15	-0.11	0.11	0.12	0.11	0.19
	weak	0.43	0.25	0.39	0.31	0.14	0.04	0.34	0.26
	strong	2.03	1.88	2.01	1.88	1.08	0.98	0.97	1.11
90%/70%	none	-0.05	-0.59	-0.06	-0.28	1.22	0.04	3.49	-0.38
	weak	0.17	0.27	0.31	0.31	1.26	0.12	4.58	-0.43
	strong	0.61	3.37	1.04	3.01	2.39	0.99	8.22	0.36
90%/90%	none	-1.05	-1.18	-0.11	-0.23	1.22	1.17	2.28	2.28
	weak	-0.27	-0.28	0.38	0.37	1.32	1.32	2.34	2.36
	strong	2.88	2.92	2.31	2.35	1.79	2.65	5.35	5.61
		True model: Plackett copula							
70%/70%	none	-0.19	-0.03	-0.17	-0.05	0.02	0.17	-0.18	0.19
	weak	1.28	0.43	1.00	1.62	0.48	3.32	-0.80	10.21
	strong	1.94	4.50	1.31	4.69	0.67	8.25	-0.87	14.47
90%/70%	none	0.04	-0.53	0.00	-0.26	1.25	0.05	3.65	-0.23
	weak	1.21	4.29	0.73	2.56	1.85	3.17	2.59	4.20
	strong	1.26	7.55	0.30	6.94	2.00	6.02	4.98	3.33
90%/90%	none	-1.11	-1.12	-0.15	-0.17	1.17	1.24	2.25	2.36
	weak	1.46	-3.79	1.48	-4.30	2.02	2.88	2.67	4.64
	strong	3.62	4.00	3.09	3.51	2.00	-14.0	4.60	7.21
		True model: Standard model							
70%/70%	none	-0.99	-0.98	-0.97	-1.06	-0.60	-0.56	-0.16	-0.53
	weak	-0.12	-0.15	-0.25	-0.12	-0.48	-0.42	-0.06	-0.19
	strong	1.42	1.33	1.50	1.41	0.70	0.57	0.33	-0.08
90%/70%	none	-0.85	-1.31	-1.08	-1.61	-0.28	-0.58	-0.37	-0.62
	weak	-0.23	-0.01	-0.42	-0.36	-0.22	-0.67	-0.30	-0.42
	strong	0.69	2.25	0.81	1.96	-0.28	-0.58	-0.01	-0.17
90%/90%	none	-1.53	-1.53	-2.11	-2.12	-0.40	-0.42	0.16	-0.09
	weak	-0.32	-0.37	-1.65	-1.66	-0.38	-0.50	-0.05	0.27
	strong	1.20	1.18	1.18	1.20	-0.22	-20.7	0.24	-0.05

Statistics
in Medicine

Table 10. Empirical coverage (in %) for the 95% confidence intervals for sensitivity and specificity on the [0,1] scale for the various combinations of the underlying true model, true values for sensitivity and specificity, true associations between sensitivity and specificity, and the four estimation methods estimation methods. Results reported here are from the situation where the number of studies in the respective meta-analysis was set to 50 and had 20 observations in the respective study group of healthy and diseased probands. Abbreviations: CG=Gauss copula, CP=Plackett copula, SG=Penalized quasi-likelihood fit assuming the standard model, SN=Gaussian quadrature assuming the standard model

True sensitivity and specificity	True association between sensitivity and specificity	Estimated model: CG Se	CG Sp	CP Se	CP Sp	SG Se	SG Sp	SN Se	SN Sp
		True model: Gauss copula							
70%/70%	none	69.4	63.5	69.6	64.0	96.7	96.2	96.0	94.1
	weak	65.5	64.7	63.9	66.5	94.3	95.1	93.7	92.4
	strong	54.6	62.4	49.9	50.4	18.5	18.9	96.8	91.7
90%/70%	none	92.6	58.6	93.2	61.0	86.5	96.8	60.1	89.5
	weak	92.8	54.0	92.6	57.3	83.8	93.9	65.8	84.4
	strong	91.6	5.9	80.4	19.0	10.1	15.5	0.0	58.1
90%/90%	none	88.8	88.7	94.5	95.2	85.7	88.1	69.3	72.4
	weak	93.8	94.0	93.6	93.6	84.0	83.4	69.5	70.4
	strong	30.4	32.1	52.1	49.7	0.9	0.4	70.7	69.4
		True model: Plackett copula							
70%/70%	none	68.2	66.7	66.4	66.2	95.3	96.0	92.1	94.8
	weak	39.3	88.5	66.7	81.8	95.0	81.0	65.5	44.8
	strong	20.9	36.9	70.2	23.7	0.6	0.1	72.3	33.3
90%/70%	none	94.2	59.5	93.8	58.9	85.6	96.5	59.8	91.0
	weak	85.5	69.5	91.1	78.9	73.6	81.9	67.2	77.2
	strong	73.0	4.1	93.3	3.1	2.4	0.9	55.2	55.6
90%/90%	none	88.3	87.9	94.6	95.1	86.2	86.3	70.6	67.9
	weak	81.6	63.9	80.2	52.5	55.8	56.3	66.6	46.1
	strong	11.5	19.4	13.6	20.4	2.4	0.0	60.5	52.9
		True model: Standard model							
70%/70%	none	5.4	5.0	6.8	13.0	96.0	97.0	79.1	87.2
	weak	7.2	11.1	11.4	13.3	97.6	97.6	94.1	88.5
	strong	28.7	30.9	19.6	23.9	1.3	1.1	81.3	71.0
90%/70%	none	7.4	4.9	4.9	4.6	75.9	73.8	81.7	72.0
	weak	9.7	8.9	13.4	9.0	41.5	40.6	82.4	83.7
	strong	4.5	29.0	13.3	22.1	75.9	73.8	66.7	70.8
90%/90%	none	4.2	0.0	0.0	1.2	9.7	9.6	91.7	100.0
	weak	5.9	2.3	4.2	0.8	3.0	3.0	100.0	92.9
	strong	13.9	9.0	8.5	6.7	41.5	0.0	100.0	100.0

Prepared using simauth.cls

Statistics in Medicine

Table 11. MSE, Mean squared error (multiplied by 100) for the estimation of sensitivity and specificity on the [0,1] scale for the various combinations of the underlying true model, true values for sensitivity and specificity, true associations between sensitivity and specificity, and the four estimation methods estimation methods. Results reported here are from the situation where the number of studies in the respective meta-analysis was set to 50 and had 20 observations in the respective study group of healthy and diseased probands. Abbreviations: CG=Gauss copula, CP=Plackett copula, SG=Penalized quasi-likelihood fit assuming the standard model, SN=Gaussian quadrature assuming the standard model

		Estimated model							
True sensitivity and specificity	True association between sensitivity and specificity	CG		CP		SG		SN	
		Se	Sp	Se	Sp	Se	Sp	Se	Sp
True model: Gauss copula									
70%/70%	none	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
	weak	0.03	0.03	0.02	0.03	0.02	0.02	0.02	0.03
	strong	0.06	0.06	0.04	0.03	0.05	0.06	0.04	0.04
90%/70%	none	0.02	0.02	0.03	0.19	0.03	0.02	0.02	0.03
	weak	0.02	0.02	0.03	0.31	0.03	0.02	0.02	0.03
	strong	0.02	0.03	0.07	0.69	0.13	0.11	0.03	0.02
90%/90%	none	0.04	0.02	0.03	0.07	0.04	0.02	0.03	0.07
	weak	0.03	0.02	0.03	0.07	0.02	0.02	0.03	0.07
	strong	0.09	0.06	0.04	0.35	0.10	0.07	0.08	0.37
True model: Plackett copula									
70%/70%	none	0.03	0.02	0.02	0.03	0.02	0.02	0.02	0.02
	weak	0.04	0.03	0.02	0.06	0.15	0.16	0.23	1.57
	strong	0.06	0.04	0.02	0.10	0.25	0.25	0.84	2.34
90%/70%	none	0.02	0.02	0.03	0.21	0.03	0.02	0.02	0.03
	weak	0.03	0.02	0.05	0.09	0.27	0.19	0.21	0.31
	strong	0.09	0.02	0.06	0.31	0.63	0.51	0.43	0.53
90%/90%	none	0.04	0.02	0.03	0.07	0.04	0.02	0.03	0.07
	weak	0.04	0.04	0.06	0.09	0.36	0.41	0.11	0.24
	strong	0.18	0.10	0.06	0.29	0.17	0.14	2.02	0.55
True model: Standard model									
70%/70%	none	0.03	0.03	0.02	0.02	0.03	0.03	0.02	0.02
	weak	0.02	0.02	0.02	0.01	0.02	0.02	0.02	0.01
	strong	0.03	0.04	0.03	0.01	0.03	0.03	0.03	0.02
90%/70%	none	0.01	0.02	0.00	0.01	0.04	0.05	0.02	0.02
	weak	0.01	0.01	0.00	0.00	0.04	0.03	0.02	0.02
	strong	0.01	0.01	0.00	0.00	0.06	0.05	0.02	0.02
90%/90%	none	0.04	0.05	0.01	0.00	0.04	0.05	0.01	0.00
	weak	0.02	0.04	0.00	0.00	0.02	0.04	0.00	0.00
	strong	0.02	0.02	0.00	0.00	0.02	0.03	4.29	0.00

Statistics
in Medicine

Table 12. Number of non-missing values (XXX/ YYY) for the estimated parameter (XXX) and the corresponding standard error (YYY) from 1000 simulation runs for the various combinations of the underlying true model, true values for sensitivity and specificity, true associations between sensitivity and specificity, and the four estimation methods estimation methods. Results reported here are from the situation where the number of studies in the respective meta-analysis was set to 50 and had 20 observations in the respective study group of healthy and diseased probands. Abbreviations: CG=Gauss copula, CP=Plackett copula, SG=Penalized quasi-likelihood fit assuming the standard model, SN=Gaussian quadrature assuming the standard model

True sensitivity and specificity	True association between sensitivity and specificity	Estimated model							
		CG		CP		SG		SN	
		Se	Sp	Se	Sp	Se	Sp	Se	Sp
		True model: Gauss copula							
70%/70%	none	632/459	632/438	652/461	652/468	1000/1000	1000/1000	174/152	174/144
	weak	626/462	626/434	655/467	655/463	1000/1000	1000/1000	96/65	96/68
	strong	707/619	707/593	698/553	698/568	211/211	211/211	80/31	80/36
90%/70%	none	799/799	799/509	841/841	841/572	1000/1000	1000/1000	300/242	300/251
	weak	826/826	826/536	839/838	839/544	999/999	999/999	208/148	208/162
	strong	576/575	576/249	756/750	756/410	166/166	166/166	95/37	95/42
90%/90%	none	1000/999	1000/999	1000/1000	1000/1000	1000/1000	1000/1000	980/980	980/980
	weak	1000/1000	1000/998	999/998	999/999	997/997	997/997	941/938	941/938
	strong	851/809	851/818	1000/1000	1000/1000	11/11	11/11	93/41	93/36
		True model: Plackett copula							
70%/70%	none	607/442	607/417	660/481	660/454	1000/1000	1000/1000	169/140	169/137
	weak	751/406	751/751	836/604	836/834	999/999	999/999	149/64	149/62
	strong	616/300	616/610	855/675	855/839	7/7	7/7	106/41	106/40
90%/70%	none	793/792	793/511	849/849	849/583	999/999	999/999	317/251	317/264
	weak	997/997	997/995	1000/1000	1000/1000	996/996	996/996	894/880	894/880
	strong	713/688	713/619	966/966	966/945	9/9	9/9	158/61	158/73
90%/90%	none	999/999	999/998	1000/1000	1000/1000	1000/1000	1000/1000	987/984	987/985
	weak	999/997	999/998	1000/1000	1000/1000	790/790	790/790	574/552	574/554
	strong	786/700	786/650	999/999	999/998	1000/1000	1000/1000	85/38	85/33
		True model: Standard model							
70%/70%	none	285/162	285/157	334/185	334/197	999/999	999/999	71/42	71/46
	weak	341/127	341/158	412/187	412/176	996/996	996/996	62/40	62/44
	strong	428/278	428/296	485/288	485/280	13/13	13/13	40/32	40/31
90%/70%	none	239/126	239/135	304/169	304/165	764/764	764/764	102/89	102/92
	weak	302/126	302/146	371/175	371/166	410/410	410/410	53/41	53/39
	strong	448/246	448/415	442/235	442/300	764/764	764/764	33/20	33/23
90%/90%	none	203/109	203/110	285/138	285/158	98/98	98/98	16/12	16/12
	weak	217/113	217/122	251/127	251/124	30/30	30/30	20/10	20/14
	strong	331/163	331/141	387/202	387/188	91/91	91/91	17/12	17/12