

An extended trivariate vine copula mixed model for meta-analysis of diagnostic studies in the presence of non-evaluable outcomes

Aristidis K. Nikoloulopoulos *

Abstract

A recent paper proposed an extended trivariate generalized linear mixed model (TGLMM) for synthesis of diagnostic test accuracy studies in the presence of non-evaluable index test results. Inspired by the aforementioned model we propose an extended trivariate vine copula mixed model that includes the TGLMM as special case, but can also operate on the original scale of sensitivity, specificity, and disease prevalence. The performance of the proposed vine copula mixed model is examined by extensive simulation studies in comparison with the TGLMM. Simulation studies showed that the TGLMM leads to biased meta-analytic estimates of sensitivity, specificity, and prevalence when the univariate random effects are misspecified. The vine copula mixed model gives nearly unbiased estimates of test accuracy indices and disease prevalence. Our general methodology is illustrated by meta-analysing coronary CT angiography studies.

Key Words: Diagnostic tests; multivariate meta-analysis; prevalence, sensitivity, specificity, summary receiver operating characteristic curves.

1 Introduction

Synthesis of diagnostic test accuracy studies is the most common medical application of multivariate meta-analysis [1–3]. The purpose of a meta-analysis of diagnostic test accuracy studies is to combine information over different studies, and provide an integrated analysis that will have more statistical power to detect an accurate diagnostic test than an analysis based on a single study.

Diagnostic test accuracy studies observe the result of a gold standard procedure that defines the presence or absence of a disease and the result of a diagnostic test. The accuracy of the diagnostic test is commonly measured by a pair of indices such as sensitivity and specificity. Sensitivity is defined

*a.nikoloulopoulos@uea.ac.uk, School of Computing Sciences, University of East Anglia, Norwich NR4 7TJ, U.K.

as the probability of testing positive given a person being diseased and specificity is defined as the probability of testing negative given a person being non-diseased [3]. The diagnostic test accuracy studies typically report the number of true positives (diseased subjects correctly diagnosed), false positives (non-diseased subjects incorrectly diagnosed as diseased), true negatives (non-diseased subjects correctly diagnosed as non-diseased) and false negatives (diseased subjects incorrectly diagnosed as non-diseased). However, diagnostic test outcomes can be non-evaluable [4]. This is the case for coronary computed tomography (CT) angiography studies which have non-evaluable results of index test in various ways such as when transferring a segment/vessel to a patient based evaluation [5].

In meta-analysis of diagnostic test accuracy studies, the existence of non-evaluable subjects is an important issue that could potentially lead to biased estimates of index test accuracy [5, 6]. Schuetz et al. [5] studied different ad-hoc approaches dealing with diagnostic test non-evaluable subjects, such as non-evaluable subjects are excluded from the study, non-evaluable positives (non-evaluable diseased subjects) are taken as true positives and non-evaluable negatives (non-evaluable non-diseased subjects) are taken as false positives, non-evaluable positives are taken as false negatives and non-evaluable negatives are taken as true negatives, and non-evaluable positives as false negatives and non-evaluable negatives as false positives. In all of these approaches, Schuetz et al. [5] used the bivariate generalized linear mixed model (BGLMM) [7], which assumes independent binomial distributions for the true positives and true negatives, conditional on the latent pair of transformed (via a link function) sensitivity and specificity in each study. They concluded that excluding the index test non-evaluable subjects leads to overestimation of the meta-analytic estimates of sensitivity and specificity and recommended the intent-to-diagnose approach by treating non-evaluable positives as false negatives and non-evaluable negatives as false positives.

Ma et al. [6] proposed a trivariate generalized linear mixed model (TGLMM) approach by treating the non-evaluable subjects as missing data under a missing at random (MAR) assumption. The TGLMM was originally proposed by Chu et al. [8] to account for potential correlations among sensitivity, specificity and disease prevalence as many empirical studies have shown the assumption of independence between the sensitivity/specificity with disease prevalence for a dichotomous disease status is likely to be violated [9–11]. Ma et al. [6] with extensive simulation studies have shown that (a) the intent-to-diagnose approach [5] under-estimates both meta-analytic estimates of sensitivity and specificity, (b) excluding the index test non-evaluable subjects does not lead to biased estimates

of sensitivity and specificity, but leads to biased estimates of prevalence, and (c) the TGLMM gives nearly unbiased estimates of the meta-analytic estimates of sensitivity, specificity and prevalence.

In this paper, inspired by Ma et al. [6], we extend the vine copula mixed model for trivariate meta-analysis of diagnostic test accuracy studies accounting for disease prevalence [12] to additionally account for non-evaluable subjects. The advantages of this methodology are that (a) the extended TGLMM is included as a special case, (b) sensitivity, specificity, and prevalence can be modelled in the original scale, and (c) tail dependencies and asymmetries can be provided.

The remainder of the paper proceeds as follows. Section 2 introduces the proposed model for diagnostic test accuracy studies in the presence of non-evaluable subjects and discusses its relationship with the extended TGLMM. Section 3 contains small-sample efficiency calculations to investigate the effect of misspecifying the random effects distribution on parameter estimates and standard errors and compares the method with the TGLMM approach. Section 4 re-evaluates the meta-analysis of coronary CT angiography studies [5, 6] using the proposed vine copula mixed model approach. We conclude with some discussion in Section 5, followed by a brief section with software details.

2 The vine copula mixed model in the presence of non-evaluable subjects

In this section, we extend the trivariate vine copula mixed model [12] to handle non-evaluable results and discuss its relationship with the extended TGLMM.

2.1 Notation

We first introduce the notation used in this paper. Let i be an index for the individual studies, j an index for the test outcome (0:negative; 1:positive; 2: non-evaluable) and k an index for the disease outcome (0: non-diseased; 1: diseased). The count data y_{ijk} , $i = 1, \dots, N$, $j = 0, 1, 2$, $k = 0, 1$, corresponding to a combination of index test j and disease outcome k in study i , form a 3×2 table (Table 1). This table has an additional row that represents the counts of non-evaluable outcomes, hence, extends the “classic” 2×2 table (Table 2).

2.2 The within-study model

For each study i , the within-study model assumes that the number of true negatives Y_{i00} , false negatives Y_{i01} , false positives Y_{i10} , true positives Y_{i11} , non-evaluable negatives Y_{i20} , and non-evaluable positives Y_{i21} are multinomially distributed given $\mathbf{X} = \mathbf{x}$, where $\mathbf{X} = (X_1, X_2, X_3, X_4, X_5)$ denotes

Table 1: Data (including the non-evaluable outcomes) from an individual study in a 3×2 table.

Test	Disease (by gold standard)		Total
	–	+	
–	y_{i00}	y_{i01}	y_{i0+}
+	y_{i10}	y_{i11}	y_{i1+}
Non-evaluable	y_{i20}	y_{i21}	y_{i2+}
Total	y_{i+0}^*	y_{i+1}^*	y_{i++}^*

Table 2: Data (excluding the non-evaluable outcomes) from an individual study in a 2×2 table.

Test	Disease (by gold standard)		Total
	–	+	
–	y_{i00}	y_{i01}	y_{i0+}
+	y_{i10}	y_{i11}	y_{i1+}
Total	y_{i+0}	y_{i+1}	y_{i++}

the transformed (via a link function $l(\cdot)$) latent vector of sensitivity, specificity, disease prevalence, probability of non-evaluable positives and probability of non-evaluable negatives, viz.

$$(Y_{i00}, Y_{i01}, Y_{i10}, Y_{i11}, Y_{i20}, Y_{i21},) | (X_1 = x_1, X_2 = x_2, X_3 = x_3, X_4 = x_4, X_5 = x_5) \sim \mathcal{M}_6(y_{i++}^*, \varpi), \ddagger$$

where

$$\begin{aligned} \varpi &= (\varpi_{00}, \varpi_{01}, \varpi_{10}, \varpi_{11}, \varpi_{20}, \varpi_{21}) \tag{1} \\ &= \left(\underbrace{l^{-1}(x_2)(1-l^{-1}(x_3))(1-l^{-1}(x_5))}_{\varpi_{00}}, \underbrace{(1-l^{-1}(x_1))l^{-1}(x_3)(1-l^{-1}(x_4))}_{\varpi_{01}}, \right. \\ &\quad \left. \underbrace{(1-l^{-1}(x_2))(1-l^{-1}(x_3))(1-l^{-1}(x_5))}_{\varpi_{10}}, \underbrace{l^{-1}(x_1)l^{-1}(x_3)(1-l^{-1}(x_4))}_{\varpi_{11}}, \right. \\ &\quad \left. \underbrace{(1-l^{-1}(x_3))l^{-1}(x_5)}_{\varpi_{20}}, \underbrace{l^{-1}(x_3)l^{-1}(x_4)}_{\varpi_{21}} \right) \end{aligned}$$

is derived by Ma et al. [6] under an MAR assumption.

The multinomial probability mass function (pmf)

$$\frac{y_{i++}^*!}{y_{i00}!y_{i01}!y_{i10}!y_{i11}!y_{i20}!y_{i21}!} \prod_{j=0}^2 \prod_{k=0}^1 \varpi_{jk}^{y_{ijk}}$$

decomposes into a product of independent binomial pmfs given the random effects, viz.

$$g(y_{i11}; y_{i+1}, l^{-1}(x_1))g(y_{i00}; y_{i+0}, l^{-1}(x_2))g(y_{i+1}^*; y_{i++}^*, l^{-1}(x_3))g(y_{i21}; y_{i+1}^*, l^{-1}(x_4))g(y_{i20}; y_{i+0}^*, l^{-1}(x_5)),$$

where

$$g(y; n, \pi) = \binom{n}{y} \pi^y (1-\pi)^{n-y}, \quad y = 0, 1, \dots, n, \quad 0 < \pi < 1,$$

$\ddagger \mathcal{M}_T(n, \mathbf{p})$ is shorthand notation for the multinomial distribution; T is the number of cells, n is the number of observations, and $\mathbf{p} = (p_1, \dots, p_T)$ with $p_1 + \dots + p_T = 1$ is the T -dimensional vector of success probabilities.

is the binomial pmf. Hence, the within-study model actually assumes that

$$\begin{aligned}
Y_{i11}|X_1 = x_1 &\sim \text{Binomial}(y_{i+1}, l^{-1}(x_1)); \\
Y_{i00}|X_2 = x_2 &\sim \text{Binomial}(y_{i+0}, l^{-1}(x_2)); \\
Y_{i+1}^*|X_3 = x_3 &\sim \text{Binomial}(y_{i++}^*, l^{-1}(x_3)),
\end{aligned} \tag{2}$$

and

$$\begin{aligned}
Y_{i21}|X_4 = x_4 &\sim \text{Binomial}(y_{i+1}^*, l^{-1}(x_4)); \\
Y_{i20}|X_5 = x_5 &\sim \text{Binomial}(y_{i+0}^*, l^{-1}(x_5)).
\end{aligned}$$

2.3 The between-studies model

Under the MAR assumption, (X_1, X_2, X_3) are independent of the missing probabilities (X_4, X_5) , hence the joint likelihood factors into two components, one involving only the transformed sensitivity x_1 , specificity x_2 and disease prevalence x_3 , and the other involving only the transformed probabilities of non-evaluable positives x_4 and non-evaluable negatives x_5 . Hence, the methodology of Nikoloulopoulos [12] can be applied to the first likelihood component to infer about the sensitivity, specificity and disease prevalence.

Nikoloulopoulos [12] proposed a vine copula mixed model as an extension of the TGLMM by rather using a vine copula representation for the random effects distribution of the latent sensitivity, specificity and disease prevalence. The trivariate vine copula can cover flexible dependence structures through the specification of 2 bivariate marginal copulas and one bivariate conditional copula that condition on 1 variable [13]. A vine requires a decision on the indexing of variables. For a 3-dimensional vine copula there are 3 distinct permutations [14]:

$$\{12, 13, 23|1\}, \quad \{12, 23, 13|2\}, \quad \text{and} \quad \{13, 23, 12|3\}.$$

To be concrete in the exposition of the theory, we use the permutation $\{12, 13, 23|1\}$; the theory though also apply to the other two permutations.

To this end, the stochastic representation of the between-studies (random effects) model takes the form

$$\left(F(X_1; l(\pi_1), \delta_1), F(X_2; l(\pi_2), \delta_2), F(X_3; l(\pi_3), \delta_3) \right) \sim C(\cdot; \boldsymbol{\theta}), \tag{3}$$

where $C(\cdot; \boldsymbol{\theta})$ is a vine copula with dependence parameter vector $\boldsymbol{\theta} = (\theta_{12}, \theta_{13}, \theta_{23|1})$ and $F(\cdot; l(\pi), \delta)$ is the cdf of the univariate distribution of the random effect. The copula parameters $\theta_{12}, \theta_{13}, \theta_{23|1}$ are

parameters of the random effects model, and they are separated from the univariate parameters. The univariate parameters π_1 , π_2 , and π_3 are those of actual interest denoting the meta-analytic parameters of sensitivity, specificity, and disease prevalence, respectively, while the univariate parameters δ_1 , δ_2 and δ_3 are of secondary interest denoting the variability between-studies for sensitivity, specificity, and disease prevalence, respectively.

The vine copula density is decomposed in a product of univariate and bivariate copula densities, viz.

$$\begin{aligned}
f_{123}(x_1, x_2, x_3; \boldsymbol{\theta}) &= f_1(x_1; l(\pi_1), \delta_1) f_2(x_2; l(\pi_2), \delta_2) f_3(x_3; l(\pi_3), \delta_3) c_{12}\left(F_1(x_1; l(\pi_1), \delta_1), \right. \\
&\quad \left. F_2(x_2; l(\pi_2), \delta_2); \theta_{12}\right) c_{13}\left(F_1(x_1; l(\pi_1), \delta_1), F_3(x_3; l(\pi_3), \delta_3); \theta_{13}\right) \times \\
&\quad c_{23|1}\left(F_{2|1}(x_2|x_1; l(\pi_1), l(\pi_2), \delta_1, \delta_2), F_{3|1}(x_3|x_1; l(\pi_1), l(\pi_3), \delta_1, \delta_3); \theta_{23|1}\right) \\
&= f_1(x_1; l(\pi_1), \delta_1) f_2(x_2; l(\pi_2), \delta_2) f_3(x_3; l(\pi_3), \delta_3) c_{123}\left(F_1(x_1; l(\pi_1), \delta_1), \right. \\
&\quad \left. F_2(x_2; l(\pi_2), \delta_2), F_3(x_3; l(\pi_3), \delta_3); \boldsymbol{\theta}\right),
\end{aligned}$$

where $f_j(\cdot; l(\pi_j), \delta_j)$ and $F_j(\cdot; l(\pi_j), \delta_j)$ are the density and cdf, respectively, of the random variable X_j , $c_{j_1j_2}(\cdot, \cdot; \theta_{j_1j_2})$ and $C_{j_1j_2}(\cdot, \cdot; \theta_{j_1j_2})$ are the bivariate copula density and cdf, respectively, for the pair of transformed variables $F_{j_1}(X_{j_1})$ and $F_{j_2}(X_{j_2})$, and $c_{23|1}(\cdot, \cdot; \theta_{23|1})$ is the bivariate copula density for the pair of transformed variables $F_{2|1}(X_2|X_1)$ and $F_{3|1}(X_3|X_1)$ where

$$F_{j_1|j_2}(x_{j_1}|x_{j_2}) = \partial C_{j_1j_2}(F_{j_1}(x_{j_1}), F_{j_2}(x_{j_2})) / \partial F_{j_2}(x_{j_2})$$

as derived in [15].

In line with our previous contributions in copula mixed models [12, 16–19] we use

- bivariate parametric copulas with different tail dependence behaviour, namely the BVN with intermediate tail dependence, Frank with tail independence, and Clayton with positive lower tail dependence. For the latter we also use its rotated versions to provide negative upper-lower tail dependence (Clayton rotated by 90°), positive upper tail dependence (Clayton rotated by 180°) and negative lower-upper tail dependence (Clayton rotated by 270°);
- the choices of $F(\cdot; l(\pi), \delta)$ and l that are given in Table 3.

Table 3: The choices of the $F(\cdot; l(\pi), \delta)$ and l in the extended trivariate vine copula mixed model.

$F(\cdot; l(\pi), \delta)$	l	π	δ
$N(\mu, \sigma)$	logit, probit, cloglog	$l^{-1}(\mu)$	σ
$\text{Beta}(\pi, \gamma)$	identity	π	γ

2.4 Likelihood and computational details for maximum likelihood estimation

For N studies the models in (2) and (3) together specify a trivariate vine copula mixed model with joint likelihood

$$L(\pi_1, \pi_2, \pi_3, \delta_1, \delta_2, \delta_3, \boldsymbol{\theta}) = \prod_{i=1}^N \int_0^1 \int_0^1 \int_0^1 g(y_{i11}; y_{i+1}, l^{-1}(x_1)) g(y_{i00}; y_{i+0}, l^{-1}(x_2)) \times g(y_{i+1}^*; y_{i++}^*, l^{-1}(x_3)) c_{123}(u_1, u_2, u_3; \boldsymbol{\theta}) du_1 du_2 du_3, \quad (4)$$

where $x_j = F^{-1}(u_j; l(\pi_j), \delta_j)$, $j = 1, 2, 3$.

Estimation of the model parameters $(\pi_1, \pi_2, \pi_3, \delta_1, \delta_2, \delta_3, \boldsymbol{\theta})$ can be approached by the standard maximum likelihood (ML) method, by maximizing the logarithm of the joint likelihood in (5). The estimated parameters can be obtained by using a quasi-Newton [20] method applied to the logarithm of the joint likelihood. This numerical method requires only the objective function, i.e., the logarithm of the joint likelihood, while the gradients are computed numerically and the Hessian matrix of the second order derivatives is updated in each iteration. The standard errors (SE) of the ML estimates can be also obtained via the gradients and the Hessian computed numerically during the maximization process.

For the vine copula mixed model numerical evaluation of the joint pmf can be achieved with the following steps:

1. Calculate Gauss-Legendre [21] quadrature points $\{u_q : q = 1, \dots, n_q\}$ and weights $\{w_q : q = 1, \dots, n_q\}$ in terms of standard uniform.
2. Convert from independent uniform random variables $\{u_{q_1} : q_1 = 1, \dots, n_{q_1}\}$, $\{u_{q_2} : q_2 = 1, \dots, n_{q_2}\}$, and $\{u_{q_3} : q_3 = 1, \dots, n_{q_3}\}$ to dependent uniform random variables v_{q_1} , $v_{q_2|q_1}$, and $v_{q_2q_3|q_1}$ that have a vine distribution $C(\cdot; \boldsymbol{\theta})$ [12]:
 - 1: Set $v_{q_1} = u_{q_1}$
 - 2: $v_{q_2|q_1} = C_{12}^{-1}(u_{q_2}|u_{q_1}; \theta_{12})$
 - 3: $t_1 = C_{23|1}^{-1}(u_{q_3}|u_{q_2}; \theta_{23|1})$

$$4: v_{q_2q_3|q_1} = C_{13}^{-1}\left(t_1|u_{q_1}; \theta_{13}\right),$$

where $C(v|u; \theta)$ and $C^{-1}(v|u; \theta)$ are the conditional copula cdf and its inverse.

3. Numerically evaluate the joint pmf

$$\int_0^1 \int_0^1 \int_0^1 g\left(y_{i11}; y_{i+1}, l^{-1}\left(F^{-1}(u_1; l(\pi_1), \delta_1)\right)\right) g\left(y_{i00}; y_{i+0}, l^{-1}\left(F^{-1}(u_2; l(\pi_2), \delta_2)\right)\right) \\ g\left(y_{i+1}^*; y_{i++}^*, l^{-1}\left(F^{-1}(u_3; l(\pi_3), \delta_3)\right)\right) c_{123}(u_1, u_2, u_3; \boldsymbol{\theta}) du_1 du_2 du_3$$

in a triple sum:

$$\sum_{q_1=1}^{n_q} \sum_{q_2=1}^{n_q} \sum_{q_3=1}^{n_q} g\left(y_{i11}; y_{i+1}, l^{-1}\left(F^{-1}(v_{q_1}; l(\pi_1), \delta_1)\right)\right) g\left(y_{i00}; y_{i+0}, l^{-1}\left(F^{-1}(v_{q_2|q_1}; l(\pi_2), \delta_2)\right)\right) \\ g\left(y_{i+1}^*; y_{i++}^*, l^{-1}\left(F^{-1}(v_{q_2q_3|q_1}; l(\pi_3), \delta_3)\right)\right).$$

With Gauss-Legendre quadrature, the same nodes and weights are used for different functions; this helps in yielding smooth numerical derivatives for numerical optimization via quasi-Newton.

2.5 Relationship with the TGLMM

In this subsection, we show what happens when all the bivariate copulas are bivariate normal (BVN) and the univariate distribution of the random effects is the $N(\mu, \sigma)$ distribution. One can easily deduce that the within-study model in (2) is the same as in the TGLMM.

Furthermore, when the three bivariate copulas are BVN copulas with copula (correlation) parameters $\rho_{12}, \rho_{13}, \rho_{23|1}$, the resulting distribution is the trivariate normal (TVN) with mean vector $\boldsymbol{\mu} = (l(\pi_1), l(\pi_2), l(\pi_3))^\top$ and variance covariance matrix $\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ \rho_{13}\sigma_1\sigma_3 & \rho_{23}\sigma_2\sigma_3 & \sigma_3^2 \end{pmatrix}$, where $\rho_{23} = \rho_{23|1}\sqrt{1 - \rho_{12}^2}\sqrt{1 - \rho_{13}^2}$. Therefore, the between-studies model in (3) assumes that $\mathbf{X} = (X_1, X_2, X_3)$ is TVN distributed, i.e., $\mathbf{X} \sim \text{TVN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

With some calculus it can be shown that the joint likelihood in (5) becomes

$$L(\pi_1, \pi_2, \pi_3, \sigma_1, \sigma_2, \sigma_3, \boldsymbol{\theta}) = \prod_{i=1}^N \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(y_{i11}; y_{i+1}, l^{-1}(x_1)) g(y_{i00}; y_{i+0}, l^{-1}(x_2)) \times \\ g(y_{i+1}^*; y_{i++}^*, l^{-1}(x_3)) \phi_{123}(x_1, x_2, x_3; \boldsymbol{\mu}, \boldsymbol{\Sigma}) dx_1 dx_2 dx_3,$$

where $\phi_{123}(\cdot; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ is the TVN density with mean vector $\boldsymbol{\mu}$ and variance covariance matrix $\boldsymbol{\Sigma}$. Hence, this model is the same as the extended TGLMM in [6].

3 Small-sample efficiency–misspecification of the random effects distribution

An extensive simulation study is conducted (a) to gauge the small-sample efficiency of the ML method, and (b) to investigate in detail the misspecification of the parametric margin or family of copulas of the random effects distribution. We also include comparisons with the TGLMM [6], that is a vine copula mixed model composed of BVN copulas and normal margins as shown in Subsection 2.5, as the current state of the art of the various meta-analytic approaches to handle non-evaluable results. We don't include either the intent-to-diagnose approach or the BGLMM that excludes the non-evaluable subjects in our simulation study as Ma et al. [6] has already established that these methods produce biased estimates in the presence of index test non-evaluable subjects.

In this simulation study we follow the configurations in Ma et al. [6]. We conduct simulation studies under three missing scenarios:

- the probabilities for non-evaluable diseased and non-diseased subjects are the same, i.e., $l^{-1}(x_4) = l^{-1}(x_5) = 0.1$;
- the probability for non-evaluable diseased subjects is smaller than the probability for non-evaluable non-diseased subjects, i.e., $l^{-1}(x_4) = 0.1 < l^{-1}(x_5) = 0.2$;
- the probability for non-evaluable non-diseased subjects is smaller than the probability for non-evaluable diseased subjects, i.e., $l^{-1}(x_4) = 0.2 > l^{-1}(x_5) = 0.1$.

All three scenarios satisfy the MAR assumption, and the first scenario also satisfies the missing completely at random assumption [22].

True sensitivity π_1 and specificity π_2 are 0.7 and 0.9, disease prevalence π_3 is 0.25 and the variability parameters are $\sigma_1 = \sigma_2 = \sigma_3 = 1$ or $\gamma_1 = \gamma_2 = \gamma_3 = 0.1$ for normal or beta margin, respectively. A moderate negative Kendall's tau association of $\tau_{12} = -0.5$ is assumed between X_1 and X_2 , a moderate positive Kendall's tau association of $\tau_{13} = 0.5$ is assumed between X_1 and X_3 , and a moderate negative Kendall's tau association of $\tau_{23|1} = -0.5$ is assumed between X_2 and X_3 given X_1 . Under each setting, 10,000 meta-analysis data sets are simulated with $N = 30$ studies in each data set. The simulation process is as below:

For $i = 1, \dots, N$:

1. Simulate (u_1, u_2, u_3) from a C-vine $C(\cdot; \tau_{12}, \tau_{13}, \tau_{23|1})$ [23]. We convert from τ 's to the BVN,

Frank and (rotated) Clayton copula parameters θ 's via the relations

$$\tau = \frac{2}{\pi} \arcsin(\theta), \quad (5)$$

$$\tau = \begin{cases} 1 - 4\theta^{-1} - 4\theta^{-2} \int_{\theta}^0 \frac{t}{e^t - 1} dt & , \quad \theta < 0 \\ 1 - 4\theta^{-1} + 4\theta^{-2} \int_0^{\theta} \frac{t}{e^t - 1} dt & , \quad \theta > 0 \end{cases}, \quad (6)$$

and

$$\tau = \begin{cases} \theta/(\theta + 2) & , \quad \text{by } 0^\circ \text{ or } 180^\circ \\ -\theta/(\theta + 2) & , \quad \text{by } 90^\circ \text{ or } 270^\circ \end{cases}, \quad (7)$$

in [24], [25], and [26], respectively.

2. Convert to beta or normal realizations via $x_j = l^{-1}\left(F_j^{-1}(u_j, l(\pi_j), \delta_j)\right)$ for $j = 1, 2, 3$.
3. Simulate the study size n from a shifted gamma distribution [27], i.e., $n \sim \text{sGamma}(\alpha = 1.2, \beta = 0.01, \text{lag} = 30)$ and round off to the nearest integer.
4. Generate $(y_{i00}, y_{i01}, y_{i10}, y_{i11}, y_{i20}, y_{i21})$ from $\mathcal{M}_6(n, \varpi)$; see (1) for the elements of the probability vector ϖ .

From the simulation results it is revealed that the maximum likelihood estimates (MLEs) are not affected by different missingness scenarios. Hence we provide here the simulation results for one missingness scenario. The results for the other two missingness scenarios are provided in the tables of the Supplementary Material. Tables 4 and 5 contain the resultant biases, standard deviations (SDs), average theoretical variances $\sqrt{\bar{V}}$, and root mean square errors (RMSEs), scaled by 100, for the MLEs under different copula choices and margins under different copula and marginal choices from the vine copula mixed model with normal (the SDs and $\sqrt{\bar{V}}$'s for the MLEs of γ_j 's instead of σ_j 's are provided for beta margins) and beta (the SDs and $\sqrt{\bar{V}}$'s for the MLEs of σ_j 's instead of γ_j 's are provided for normal margins) margins, respectively. The theoretical variances of the MLEs are obtained via the gradients and the Hessian that were computed numerically during the maximization process. The true (simulated) copula distributions are the Clayton copulas rotated by 90° for both the $C_{12}(\cdot; \tau_{12})$ and $C_{13}(\cdot; \tau_{23|1})$ copulas and the Clayton copula for the $C_{13}(\cdot; \tau_{13})$ copula.

Conclusions from the values in the tables are the following:

- ML with the true vine copula mixed model is highly efficient according to the simulated biases, SDs and RMSEs.
- The MLEs of π_j , $j = 1, 2, 3$ are not robust to margin misspecification, e.g., in Table 4 (Table 5) where the true univariate margins are normal (beta) the scaled biases for the MLEs of π_1 for

Table 4: Small sample of sizes $N = 30$ simulations (10^4 replications, $n_q = 15$) from the extended trivariate vine copula mixed model with normal margins and biases, root mean square errors (RMSEs) and standard deviations (SDs), along with the square root of the average theoretical variances ($\sqrt{\bar{V}}$), scaled by 100, for the MLEs under different copula choices and margins. The true (simulated) copula distributions are the Clayton copulas rotated by 90° for both the $C_{12}(\cdot; \tau_{12})$ and $C_{13}(\cdot; \tau_{23|1})$ copulas and the Clayton copula for the $C_{13}(\cdot; \tau_{13})$ copula. The missing probability of diseased group is smaller than non-diseased group, i.e., $l^{-1}(x_4) = 0.1 < l^{-1}(x_5) = 0.2$. The SDs and $\sqrt{\bar{V}}$'s for the MLEs of γ_j 's instead of σ_j 's are provided for beta margins.

	margin	copula	π_1	π_2	π_3	σ_1	σ_2	σ_3	τ_{12}	τ_{13}	$\tau_{23 1}$
Bias	† normal	BVN	0.19	-0.12	0.11	-7.96	-2.36	-1.90	-4.87	-3.33	16.76
		Frank	0.22	-0.16	-0.27	-7.44	-1.47	-2.31	-6.24	-1.14	15.21
	§	$\text{Cln}\{0^\circ, 90^\circ\}$	0.32	-0.04	0.88	-6.96	1.92	-3.97	-3.73	-6.85	20.71
		$\text{Cln}\{0^\circ, 270^\circ\}$	1.27	-0.22	0.01	-9.14	-1.80	1.15	4.94	-17.02	16.67
	beta	BVN	-2.07	-3.57	4.04	-	-	-	-3.93	-4.68	18.46
		Frank	-1.76	-3.56	3.55	-	-	-	-5.90	-1.81	17.34
		$\text{Cln}\{0^\circ, 90^\circ\}$	-1.82	-3.71	4.48	-	-	-	-2.78	-9.76	19.03
		$\text{Cln}\{0^\circ, 270^\circ\}$	-1.14	-3.72	4.15	-	-	-	5.39	-17.35	19.60
SD	† normal	BVN	4.37	1.86	3.55	17.84	16.56	14.73	15.38	13.20	23.98
		Frank	4.61	1.93	3.60	18.59	17.28	14.81	16.66	14.04	25.02
	§	$\text{Cln}\{0^\circ, 90^\circ\}$	4.53	1.91	3.79	19.66	18.88	15.64	14.28	17.85	24.96
		$\text{Cln}\{0^\circ, 270^\circ\}$	4.28	2.00	3.71	18.96	18.48	17.35	27.31	17.87	22.77
	beta	BVN	3.95	2.27	3.37	3.97	3.03	3.35	15.08	13.19	23.61
		Frank	4.13	2.34	3.38	4.14	3.07	3.34	16.45	14.03	25.55
		$\text{Cln}\{0^\circ, 90^\circ\}$	4.10	2.43	3.67	4.35	3.69	3.53	14.52	17.58	26.02
		$\text{Cln}\{0^\circ, 270^\circ\}$	3.91	2.41	3.58	4.21	3.19	4.13	26.97	18.78	22.80
$\sqrt{\bar{V}}$	† normal	BVN	3.95	1.77	3.36	16.76	15.84	13.52	12.73	11.05	18.65
		Frank	3.90	1.74	3.20	17.04	16.13	13.19	12.38	10.96	17.46
	§	$\text{Cln}\{0^\circ, 90^\circ\}$	3.73	1.69	2.83	16.04	14.72	10.77	11.20	8.37	13.52
		$\text{Cln}\{0^\circ, 270^\circ\}$	3.42	1.65	3.02	15.86	14.61	13.07	11.66	7.72	10.44
	beta	BVN	3.54	2.00	3.04	3.78	2.54	3.02	12.79	11.22	18.70
		Frank	3.51	1.94	2.89	3.87	2.54	2.88	12.68	11.04	17.78
		$\text{Cln}\{0^\circ, 90^\circ\}$	3.41	1.83	2.48	3.58	2.46	2.30	11.57	8.79	13.99
		$\text{Cln}\{0^\circ, 270^\circ\}$	3.13	1.83	2.70	3.53	2.31	2.82	12.26	8.12	10.35
RMSE	† normal	BVN	4.37	1.86	3.55	19.54	16.72	14.85	16.13	13.61	29.26
		Frank	4.62	1.93	3.61	20.02	17.34	14.99	17.79	14.09	29.28
	§	$\text{Cln}\{0^\circ, 90^\circ\}$	4.54	1.91	3.89	20.86	18.98	16.14	14.76	19.11	32.43
		$\text{Cln}\{0^\circ, 270^\circ\}$	4.46	2.01	3.71	21.05	18.56	17.39	27.75	24.67	28.22
	beta	BVN	4.46	4.24	5.26	-	-	-	15.58	14.00	29.97
		Frank	4.49	4.26	4.90	-	-	-	17.48	14.15	30.88
		$\text{Cln}\{0^\circ, 90^\circ\}$	4.49	4.44	5.79	-	-	-	14.78	20.11	32.24
		$\text{Cln}\{0^\circ, 270^\circ\}$	4.07	4.43	5.48	-	-	-	27.50	25.57	30.06

§: True model; †: The resulting model is the same as the TGLMM; $\text{Cln}\{\omega_1^\circ, \omega_2^\circ\}$: The $C_{13}(\cdot; \tau_{13})$ and $\{C_{12}(\cdot; \tau_{12}), C_{23|1}(\cdot; \tau_{23|1})\}$ pair copulas are Clayton rotated by ω_1 and ω_2 degrees, respectively.

the various vine copula mixed models with beta (normal) margins range from -2.07 (2.10) to -1.14 (2.96).

- The SDs of π_j , $j = 1, 2, 3$ are not robust to margin misspecification, e.g., in Table 4 where the

true univariate margins are normal the scaled SDs for the MLEs of π_1 for the various extended vine copula mixed models with normal margins range from 4.28 to 4.61, while for the various extended vine copula mixed models with beta margins range from 3.91 to 4.13.

- The MLEs of π_j , $j = 1, 2, 3$ are rather robust to bivariate copula misspecification, but their biases increase when the assumed bivariate copula has tail dependence of opposite direction from the true bivariate copula. For example, in Table 4 (Table 5) the scaled biases for the MLEs of π_1 for the various vine copula mixed models with normal (beta) margins range from 0.19 (0.64) to 0.32 (0.71), but the scaled bias rises to 1.27 (1.34) when the Clayton copula rotated by 270° is assumed for both the $C_{12}(\cdot; \tau_{12})$ and $C_{13}(\cdot; \tau_{23|1})$ copulas.
- The SDs of π_j , $j = 1, 2, 3$ are robust to bivariate copula misspecification, e.g., in Table 4 (Table 5) the scaled SDs for the MLEs of π_1 for the various vine copula mixed models with normal (beta) margins ranges from 4.28 (3.21) to 4.61 (3.38).
- The MLEs of σ_j or γ_j , $j = 1, 2, 3$ are rather robust to bivariate copula misspecification, but their biases increase when the assumed bivariate copula has tail dependence of opposite direction from the true bivariate copula. For example, in Table 4 (Table 5) the scaled biases for the MLEs of σ_1 (γ_1) for the various vine copula mixed models with normal (beta) margins range from -7.96 (-1.54) to -6.96 (-1.50), but the scaled bias drops to -9.14 (-1.91) when the Clayton copula rotated by 270° is called for both the $C_{12}(\cdot; \tau_{12})$ and $C_{13}(\cdot; \tau_{23|1})$ copulas.
- The SDs of σ_j or γ_j are robust to bivariate copula misspecification, e.g. in Table 4 the scaled SDs of σ_1 range from 17.84 to 18.96 and in Table 5 the scaled SDs of γ_1 range from 3.16 to 3.35.
- The ML estimates of τ 's are robust to margin misspecification, as the copula remains invariant under any series of strictly increasing transformations of the components of the random vector, e.g., in Table 5 the scaled bias of $\hat{\tau}_{12}$ is -4.13 for the true extended vine copula mixed model and -4.18 for an extended vine copula mixed model with the true bivariate copulas but normal margins.
- The SDs of τ 's are robust to margin misspecification (for the same reason as above), e.g., in Table 5 the scaled SD of $\hat{\tau}_{12}$ is 15.01 for the true extended vine copula mixed model and 14.82

for an extended vine copula mixed model with the true bivariate copulas but normal margins.

Table 5: Small sample of sizes $N = 30$ simulations (10^4 replications, $n_q = 15$) from the extended trivariate vine copula mixed model with beta margins and biases, root mean square errors (RMSEs) and standard deviations (SDs), along with the square root of the average theoretical variances ($\sqrt{\bar{V}}$), scaled by 100, for the MLEs under different copula choices and margins. The true (simulated) copula distributions are the Clayton copulas rotated by 90° for both the $C_{12}(\cdot; \tau_{12})$ and $C_{13}(\cdot; \tau_{23|1})$ copulas and the Clayton copula for the $C_{13}(\cdot; \tau_{13})$ copula. The missing probability of diseased group is smaller than non-diseased group, i.e., $l^{-1}(x_4) = 0.1 < l^{-1}(x_5) = 0.2$. The SDs and $\sqrt{\bar{V}}$'s for the MLEs of σ_j 's instead of γ_j 's are provided for normal margins.

	margin	copula	π_1	π_2	π_3	γ_1	γ_2	γ_3	τ_{12}	τ_{13}	$\tau_{23 1}$	
Bias	† normal	BVN	2.25	3.59	-2.79	-	-	-	-6.31	-2.51	14.22	
		Frank	2.10	3.53	-2.92	-	-	-	-7.57	-0.71	11.71	
		$\text{Cln}\{0^\circ, 90^\circ\}$	2.16	3.68	-2.54	-	-	-	-4.18	-2.36	15.41	
		$\text{Cln}\{0^\circ, 270^\circ\}$	2.96	3.51	-2.80	-	-	-	0.28	-12.14	10.29	
beta	§	BVN	0.64	-0.02	0.10	-1.51	-0.37	-0.57	-6.72	-3.62	17.75	
		Frank	0.71	-0.11	-0.13	-1.50	-0.19	-0.68	-8.26	-1.04	16.53	
		$\text{Cln}\{0^\circ, 90^\circ\}$	0.71	-0.01	0.19	-1.54	-0.03	-1.03	-4.13	-5.44	14.81	
		$\text{Cln}\{0^\circ, 270^\circ\}$	1.34	-0.22	0.20	-1.91	-0.23	-0.17	-0.65	-12.45	14.28	
SD	† normal	BVN	3.50	1.60	2.77	16.20	21.16	13.06	17.23	15.58	30.77	
		Frank	3.68	1.66	2.80	16.59	22.27	13.15	18.74	16.76	31.27	
		$\text{Cln}\{0^\circ, 90^\circ\}$	3.66	1.65	2.91	17.36	22.00	13.87	14.82	20.92	28.61	
		$\text{Cln}\{0^\circ, 270^\circ\}$	3.44	1.71	2.85	17.04	23.62	14.94	28.32	20.78	29.77	
	beta	§	BVN	3.25	1.86	2.60	3.16	2.97	2.42	17.38	15.53	31.40
			Frank	3.38	1.96	2.61	3.23	3.13	2.39	19.16	16.66	32.36
	beta	§	$\text{Cln}\{0^\circ, 90^\circ\}$	3.38	1.94	2.74	3.35	3.26	2.44	15.01	20.93	29.30
			$\text{Cln}\{0^\circ, 270^\circ\}$	3.21	2.01	2.70	3.27	3.26	2.81	29.16	22.00	29.70
$\sqrt{\bar{V}}$	† normal	BVN	3.25	1.43	2.63	14.90	20.36	11.73	14.34	13.37	24.48	
		Frank	3.21	1.43	2.53	14.87	20.93	11.60	13.08	13.11	21.31	
		$\text{Cln}\{0^\circ, 90^\circ\}$	3.14	1.35	2.34	14.45	18.53	10.01	11.83	10.51	16.45	
		$\text{Cln}\{0^\circ, 270^\circ\}$	2.89	1.37	2.51	14.12	19.35	11.70	11.18	10.26	13.86	
	beta	§	BVN	3.05	1.80	2.48	3.01	2.84	2.30	14.61	13.44	24.71
			Frank	3.04	1.79	2.40	3.02	2.93	2.24	13.96	13.34	22.40
	beta	§	$\text{Cln}\{0^\circ, 90^\circ\}$	3.00	1.63	2.16	2.88	2.60	1.88	12.54	11.10	17.33
			$\text{Cln}\{0^\circ, 270^\circ\}$	2.78	1.71	2.40	2.80	2.73	2.29	12.37	10.98	14.92
RMSE	† normal	BVN	4.16	3.93	3.93	-	-	-	18.35	15.78	33.90	
		Frank	4.24	3.90	4.05	-	-	-	20.21	16.77	33.39	
		$\text{Cln}\{0^\circ, 90^\circ\}$	4.25	4.03	3.87	-	-	-	15.39	21.05	32.50	
		$\text{Cln}\{0^\circ, 270^\circ\}$	4.53	3.90	4.00	-	-	-	28.32	24.07	31.50	
	beta	§	BVN	3.31	1.86	2.60	3.50	3.00	2.49	18.63	15.94	36.07
			Frank	3.46	1.96	2.62	3.56	3.13	2.49	20.87	16.69	36.33
	beta	§	$\text{Cln}\{0^\circ, 90^\circ\}$	3.45	1.94	2.75	3.69	3.26	2.65	15.56	21.62	32.83
			$\text{Cln}\{0^\circ, 270^\circ\}$	3.48	2.03	2.71	3.79	3.27	2.82	29.17	25.27	32.95

§: True model; †: The resulting model is the same as the TGLMM; $\text{Cln}\{\omega_1^\circ, \omega_2^\circ\}$: The $C_{13}(\cdot; \tau_{13})$ and $\{C_{12}(\cdot; \tau_{12}), C_{23|1}(\cdot; \tau_{23|1})\}$ pair copulas are Clayton rotated by ω_1 and ω_2 degrees, respectively.

These results are in line with our previous studies [12, 16–19]. The estimation of the univariate meta-analytic parameters is a univariate inference, and hence it is the univariate marginal distribution

that matters and not the type of the copula. The extended TGLMM [6] is restricted to normal margins, hence as shown in Table 5 and Supplementary Tables 3 and 4, leads to biased estimates of the meta-analytic parameters of sensitivity, specificity and prevalence when the true univariate distribution of the latent sensitivity, specificity, and disease prevalence is beta. This is due to the fact that the vine copula mixed models with beta margins operate on the original scale of sensitivity, specificity and disease prevalence, which is not the case for the TGLMM that solely operates on a transformed scale.

4 Re-evaluation of the meta-analysis of coronary CT angiography studies

We illustrate the use of the vine copula mixed model for the meta-analysis of diagnostic accuracy studies in the presence of non-evaluable subjects by re-analysing the data on 26 studies from a systematic review for diagnostic accuracy studies of coronary CT angiography [5, 6].

We fit the vine copula mixed model for all different permutations, choices of parametric families of copulas and margins. To make it easier to compare strengths of dependence, we convert from τ to the BVN, Frank and (rotated) Clayton copula parameter θ via the relations in (5), (6), and (7). Since the number of parameters is the same between the models, we use the maximized log-likelihood that corresponds to the estimates as a rough diagnostic measure for goodness of fit between the models.

In Table 6 we present the results from the first permutation, as a different indexing didn't lead to any significant differences due to the small sample size. This is consistent with our previous study on trivariate vine copula mixed models [12]. The log-likelihoods showed that a vine copula mixed model with the Clayton copula for the $C_{12}(\cdot; \tau_{12})$ copula and the Clayton copula rotated by 90° for both the $C_{13}(\cdot; \tau_{13})$ and $C_{23|1}(\cdot; \tau_{23|1})$ copulas and beta margins provides the best fit (Table 6). It is also revealed that a vine copula mixed model with the sensitivity, specificity, and prevalence on the original scale provides better fit than the TGLMM, which models the sensitivity, specificity and prevalence on a transformed scale.

Though typically the focus of meta-analysis has been to derive the summary-effect estimates, there is increasing interest in drawing predictive inference. A summary receiver operating characteristic (SROC) curve has been deduced for the bivariate copula mixed model [16] through a median regression curve of X_1 on X_2 . However, as there is no priori reason to regress X_1 on X_2 instead of the other way around, Nikoloulopoulos [16] has also provided a median regression curve of X_2 on X_1 . In addition to using just median regression curves, quantile regression curves with a focus on high

Table 6: Maximised log-likelihoods, ML estimates and standard errors (SE) of the trivariate vine copula mixed models for diagnostic accuracy studies of coronary CT angiography.

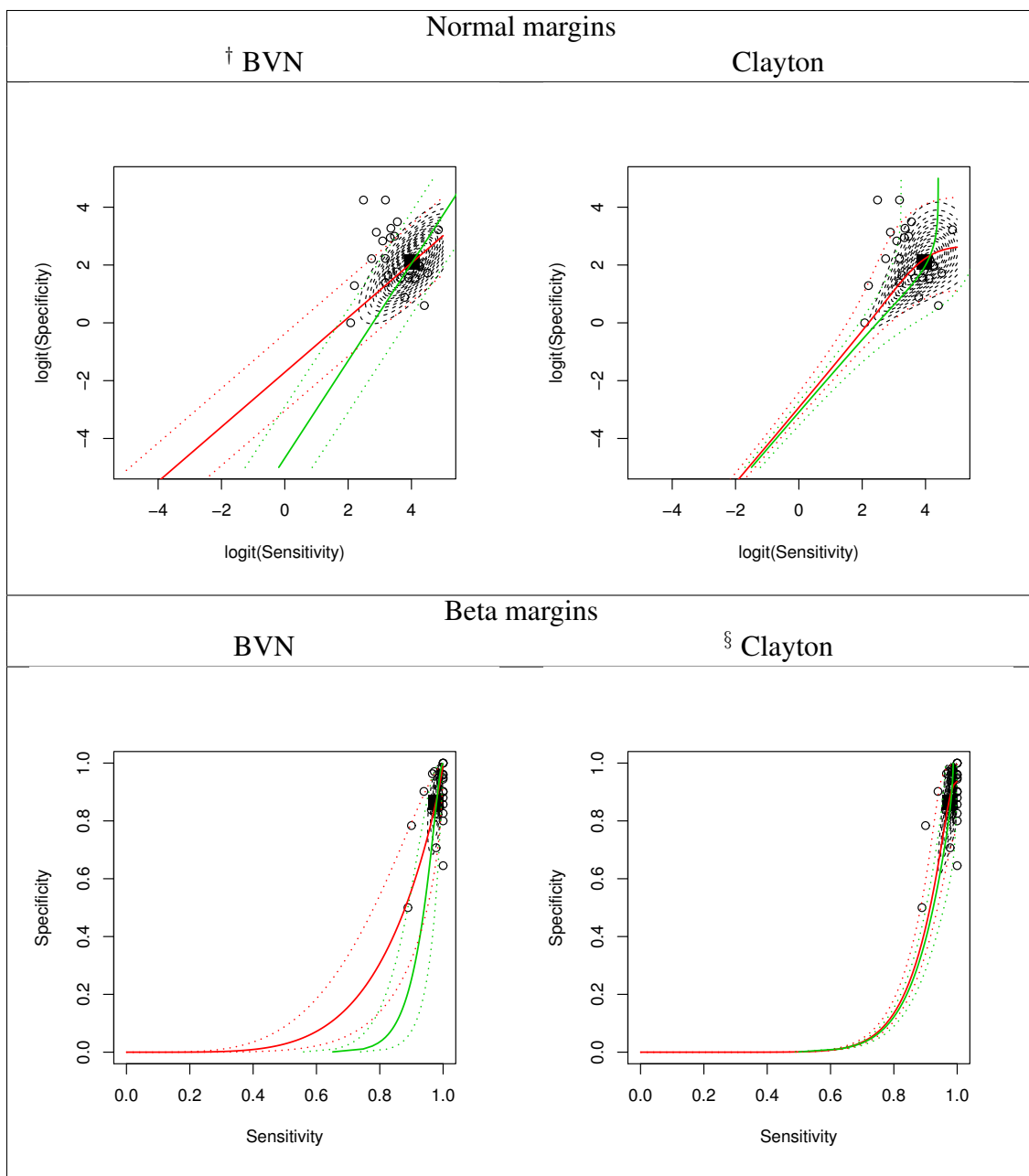
	Normal margins							
	BVN †		Cln{0°, 90°}		Cln{0°, 270°}		Frank	
	Est.	SE	Est.	SE	Est.	SE	Est.	SE
π_1	0.982	0.006	0.982	0.006	0.982	0.006	0.980	0.005
π_2	0.890	0.021	0.892	0.021	0.891	0.021	0.885	0.022
π_3	0.481	0.040	0.482	0.039	0.484	0.039	0.478	0.039
σ_1	0.687	0.343	0.670	0.347	0.684	0.328	0.478	0.291
σ_2	0.866	0.200	0.863	0.198	0.843	0.189	0.878	0.190
σ_3	0.790	0.115	0.781	0.094	0.808	0.104	0.753	0.118
τ_{12}	0.539	0.374	0.391	0.375	0.439	0.364	0.815	0.203
τ_{13}	-0.110	0.227	0.018	0.346	-0.058	0.108	-0.026	0.260
$\tau_{23 1}$	-0.231	0.312	-0.320	0.281	-0.040	0.128	-0.911	0.132
$\log L$	-194.9		-194.3		-195.4		-194.4	
	Beta margins							
	BVN		Cln{0°, 90°} §		Cln{0°, 270°}		Frank	
	Est.	SE	Est.	SE	Est.	SE	Est.	SE
π_1	0.978	0.006	0.977	0.006	0.978	0.005	0.977	0.005
π_2	0.864	0.022	0.865	0.022	0.865	0.021	0.856	0.023
π_3	0.484	0.034	0.483	0.032	0.487	0.032	0.480	0.034
γ_1	0.010	0.011	0.009	0.011	0.009	0.011	0.005	0.006
γ_2	0.076	0.031	0.075	0.031	0.073	0.029	0.081	0.031
γ_3	0.118	0.027	0.115	0.022	0.123	0.025	0.110	0.028
τ_{12}	0.519	0.367	0.407	0.456	0.465	0.490	0.797	0.205
τ_{13}	-0.105	0.225	0.033	0.272	-0.057	0.107	-0.016	0.252
$\tau_{23 1}$	-0.241	0.282	-0.345	0.234	-0.040	0.124	-0.911	0.104
$\log L$	-194.5		-193.9		-195.2		-194.0	

§: Best fit; †: The resulting model is the same as the TGLMM; Cln{ $\omega_1^\circ, \omega_2^\circ$ }: The $C_{12}(\cdot; \tau_{12})$ and $\{C_{13}(\cdot; \tau_{13}), C_{23|1}(\cdot; \tau_{23|1})\}$ pair copulas are Clayton rotated by ω_1 and ω_2 degrees, respectively.

($q = 0.99$) and low quantiles ($q = 0.01$), which are strongly associated with the upper and lower tail dependence imposed from each parametric family of copulas, have also been proposed [16]. These can be seen as confidence regions of the median regression SROC curve. Finally, a contour plot of the the random effects distribution at the ML estimate has been proposed to preserve the nature of a bivariate response instead of a univariate response along with a covariate [16]. The contour plot can be seen as the predictive region of the estimated pair of sensitivity and specificity. The prediction region of the copula mixed model does not depend on the assumption of bivariate normality of the random effects as in the TGLMM and therefore has a non-elliptical shape.

Figure 1 demonstrates the SROC curves and summary operating points (a pair of average sensitivity and specificity) with a confidence and a predictive region from the best fitted and BVN copula

Figure 1: Contour plots (predictive region) and quantile regression curves from the best fitted and BVN copula with normal (upper panel graph) and beta (lower panel graph) margins. For normal margins, the axes are in logit scale since we also plot the estimated contour plot of the random effects distribution as predictive region; this has been estimated for the logit pair of (Sensitivity, Specificity).



§: Best fit; †: The resulting model is the same as the TGLMM; ■: summary point; ○: study estimate; Red and green lines represent the quantile regression curves $x_1 := \tilde{x}_1(x_2, q)$ and $x_2 := \tilde{x}_2(x_1, q)$, respectively; for $q = 0.5$ solid lines and for $q \in \{0.01, 0.99\}$ dotted lines (confidence region).

with normal (upper panel graph) and beta (lower panel graph) margins. From the graph it is apparent that better prediction is achieved when a Clayton copula with beta margins is assumed for the random effects distribution of the latent sensitivity and specificity.

5 Discussion

We have exemplified the vine copula mixed model for trivariate meta-analysis of diagnostic test accuracy studies [12] in the presence of non-evaluable subjects. It includes the extended TGLMM [6] as a special case and it can be seen to provide an improvement over the latter on the basis of the log-likelihood principle. Hence, superior statistical inference for the meta-analytic parameters of interest can be achieved when there is a belief in an MAR assumption.

This improvement relies on the fact that the random effects distribution is expressed via vine copulas. The TVN distribution of the transformed latent proportions in the TGLMM has restricted properties, i.e., a linear correlation structure and normal margins. Copulas break the model building process into two separate steps, the choice of arbitrary marginal distributions, and the choice of an arbitrary copula function (dependence structure). Hence, we can use beta instead of normal margins to model the latent proportions in the original scale. The choice of the copula couldn't be other than the class of vine copulas. Vine copulas allow for flexible tail dependence, different from assuming simple linear correlation structures, tail independence and normality [13], which makes them well suited for meta-analysis of diagnostic tests as the traditional assumption of multivariate normality is invalid.

It has been reported in the literature that in the TGLMM estimation problems relating to the correlation parameters exist, such as non-convergence and a singular covariance matrix, particularly if the sample size is small [8]. Nevertheless, we rather propose a numerically stable ML estimation technique based on Gauss-Legendre quadrature; the crucial step is to convert from independent to dependent quadrature points. The application example and simulations use a sufficient number of individual studies, i.e., $N = 30$. For meta-analyses with fewer studies one can simplify the model using a truncated at level 1 vine copula. We refer the interested reader to our previous study on trivariate vine copula mixed models [12] for simulations and various application examples that involve a small number of studies and call this notion of a truncated at level 1 vine copula. The improvement over the reduction of the dependence parameters is small (one dependence parameter less), but for estimation purposes this is extremely useful for a small number of studies.

In an era of evidence-based medicine, decision makers need high-quality procedures such as the SROC curves to support decisions about whether or not to use a diagnostic test in a specific clinical situation. Different SROC curves essentially show the effect of different model (random effect distri-

bution) assumptions, since SROC is an inference that depends on the joint distribution. For the vine copula mixed model, the model parameters (including dependence parameters), the choice of the pair copulas, and the choice of the margin affect the shape of the SROC curve [16], while the SROC curve from the TGLMM is severely restricted to the elliptical (linear) shape.

A recurrent theme underlying our methodology for analysis in the presence of missing data is the need to make assumptions that cannot be verified based on the observed data. Throughout this paper we adopted the assumption of MAR. Nevertheless, it is natural to be concerned about robustness or sensitivity of inferences to departures from the MAR assumption. Future research will focus to handle the case when the non-evaluable subjects will be treated as non-missing categories.

Software

R functions to derive estimates and simulate from the vine copula mixed model for trivariate meta-analysis of diagnostic studies in the presence of non-evaluable subjects are part of the R package `CopulaREMADA` [28]. The data and code used in Section 4 are given as code examples in the package.

Acknowledgements

The simulations presented in this paper were carried out on the High Performance Computing Cluster supported by the Research and Specialist Computing Support service at the University of East Anglia.

References

- [1] D. Jackson, R. Riley, and I. R. White. Multivariate meta-analysis: Potential and promise. *Statistics in Medicine*, 30(20):2481–2498, 2011.
- [2] D. Mavridis and G. Salanti. A practical introduction to multivariate meta-analysis. *Statistical Methods in Medical Research*, 22(2):133–158, 2013.
- [3] X. Ma, L. Nie, S. R Cole, and H. Chu. Statistical methods for multivariate meta-analysis of diagnostic tests: An overview and tutorial. *Statistical Methods in Medical Research*, 25:1596–1619, 2016.
- [4] C. B. Begg, R. A. Greenes, and B. Iglewicz. The influence of uninterpretability on the assessment of diagnostic tests. *Journal of Chronic Diseases*, 39(8):575–584, 1986.

- [5] G. M. Schuetz, P. Schlattmann, and M. Dewey. Use of 3×2 tables with an intention to diagnose approach to assess clinical performance of diagnostic tests: Meta-analytical evaluation of coronary CT angiography studies. *BMJ (Online)*, 345:e6717, 2012.
- [6] X. Ma, M. F. K. Suri, and H. Chu. A trivariate meta-analysis of diagnostic studies accounting for prevalence and non-evaluable subjects: Re-evaluation of the meta-analysis of coronary CT angiography studies. *BMC Medical Research Methodology*, 14:128, 2014.
- [7] H. Chu and S. R. Cole. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. *Journal of Clinical Epidemiology*, 59(12):1331–1332, 2006.
- [8] H. Chu, L. Nie, S. R. Cole, and C. Poole. Meta-analysis of diagnostic accuracy studies accounting for disease prevalence: Alternative parameterizations and model selection. *Statistics in Medicine*, 28(18):2384–2399, 2009.
- [9] H. Brenner and O. Gefeller. Variation of sensitivity, specificity, likelihood ratios and predictive values with disease prevalence. *Statistics in Medicine*, 16(9):981–991, 1997.
- [10] M. M. G. Leeflang, P. M. M. Bossuyt, and L. Irwig. Diagnostic test accuracy may vary with prevalence: implications for evidence-based diagnosis. *Journal of Clinical Epidemiology*, 62(1):5–12, 2009.
- [11] M. M. G. Leeflang, A. W. S. Rutjes, J. B. Reitsma, L. Hooft, and P. M. M. Bossuyt. Variation of a test’s sensitivity and specificity with disease prevalence. *Canadian Medical Association Journal*, 185(11):E537–E544, 2013.
- [12] A. K. Nikoloulopoulos. A vine copula mixed effect model for trivariate meta-analysis of diagnostic test accuracy studies accounting for disease prevalence. *Statistical Methods in Medical Research*, 26(5):2270–2286, 2017.
- [13] H. Joe, H. Li, and A. K. Nikoloulopoulos. Tail dependence functions and vine copulas. *Journal of Multivariate Analysis*, 101:252–270, 2010.
- [14] K. Aas, C. Czado, A. Frigessi, and H. Bakken. Pair-copula constructions of multiple dependence. *Insurance: Mathematics & Economics*, 44:182–198, 2009.
- [15] H. Joe. Families of m -variate distributions with given margins and $m(m-1)/2$ bivariate dependence parameters. In L. Rüschendorf, B. Schweizer, and M.D. Taylor, editors, *Distributions with Fixed Marginals and Related Topics*, volume 28, pages 120–141, Hayward, CA, 1996. Institute of Mathematical Statistics.

- [16] A. K. Nikoloulopoulos. A mixed effect model for bivariate meta-analysis of diagnostic test accuracy studies using a copula representation of the random effects distribution. *Statistics in Medicine*, 34:3842–3865, 2015.
- [17] A. K. Nikoloulopoulos. Hybrid copula mixed models for combining case-control and cohort studies in meta-analysis of diagnostic tests. *Statistical Methods in Medical Research*, 27(8): 2540–2553, 2018.
- [18] A. K. Nikoloulopoulos. On composite likelihood in bivariate meta-analysis of diagnostic test accuracy studies. *AStA Advances in Statistical Analysis*, 102:211–227, 2018.
- [19] A. K. Nikoloulopoulos. A D-vine copula mixed model for joint meta-analysis and comparison of diagnostic tests. *Statistical Methods in Medical Research*, 28(10-11):3286–3300, 2019.
- [20] J.C. Nash. *Compact Numerical Methods for Computers: Linear Algebra and Function Minimization*. Hilger, New York, 1990. 2nd edition.
- [21] A. H. Stroud and D. Secrest. *Gaussian Quadrature Formulas*. Prentice-Hall, Englewood Cliffs, NJ, 1966.
- [22] R. J. Little and D. Rubin. *Statistical Analysis with Missing Data*. John Wiley & Sons, 2002.
- [23] H. Joe. Dependence comparisons of vine copulae with four or more variables. In D. Kurowicka and H. Joe, editors, *Dependence Modeling: Handbook on Vine Copulae*, chapter 7, pages 139–164. World Scientific, Singapore, 2011.
- [24] H. Hult and F. Lindskog. Multivariate extremes, aggregation and dependence in elliptical distributions. *Advances in Applied Probability*, 34:587–608, 2002.
- [25] C. Genest. Frank’s family of bivariate distributions. *Biometrika*, 74(3):549–555, 1987. ISSN 0006-3444.
- [26] C. Genest and J. MacKay. The joy of copulas: bivariate distributions with uniform marginals. *The American Statistician*, 40(4):280–283, 1986. ISSN 0003-1305.
- [27] M. Paul, A. Riebler, L. M. Bachmann, H. Rue, and L. Held. Bayesian bivariate meta-analysis of diagnostic test studies using integrated nested laplace approximations. *Statistics in Medicine*, 29(12):1325–1339, 2010.
- [28] A. K. Nikoloulopoulos. *CopulaREMADA: Copula mixed models for multivariate meta-analysis of diagnostic test accuracy studies*, 2019. R package version 1.3. URL: <http://CRAN.R-project.org/package=CopulaREMADA>.

Supplementary Material for

An extended trivariate vine copula mixed model for meta-analysis of diagnostic studies in the presence of non-evaluable outcomes

Aristidis K. Nikoloulopoulos *

Table 1: Small sample of sizes $N = 30$ simulations (10^4 replications, $n_q = 15$) from the extended trivariate vine copula mixed model with normal margins and biases, root mean square errors (RMSEs) and standard deviations (SDs), along with the square root of the average theoretical variances ($\sqrt{\bar{V}}$), scaled by 100, for the MLEs under different copula choices and margins. The true (simulated) copula distributions are the Clayton copulas rotated by 90° for both the $C_{12}(\cdot; \tau_{12})$ and $C_{13}(\cdot; \tau_{23|1})$ copulas and the Clayton copula for the $C_{13}(\cdot; \tau_{13})$ copula. The missing probabilities for diseased and non-diseased subjects are the same, i.e., $l^{-1}(x_4) = l^{-1}(x_5) = 0.1$. The SDs and $\sqrt{\bar{V}}$'s for the MLEs of γ_j 's instead of σ_j 's are provided for beta margins.

	margin	copula	π_1	π_2	π_3	σ_1	σ_2	σ_3	τ_{12}	τ_{13}	$\tau_{23 1}$
Bias	† normal	BVN	0.16	-0.12	0.10	-7.85	-2.23	-1.73	-4.67	-3.19	17.21
		Frank	0.20	-0.16	-0.27	-7.31	-1.24	-2.13	-5.95	-0.96	15.46
	§	$\text{Cln}\{0^\circ, 90^\circ\}$	0.30	-0.05	0.90	-6.62	2.00	-3.94	-3.64	-6.93	20.98
		$\text{Cln}\{0^\circ, 270^\circ\}$	1.25	-0.24	0.04	-9.21	-1.58	1.69	5.31	-16.72	17.55
	beta	BVN	-2.09	-3.59	4.05	-	-	-	-4.02	-4.53	19.03
		Frank	-1.80	-3.56	3.55	-	-	-	-5.86	-1.62	17.51
		$\text{Cln}\{0^\circ, 90^\circ\}$	-1.82	-3.76	4.52	-	-	-	-2.67	-10.09	19.11
		$\text{Cln}\{0^\circ, 270^\circ\}$	-1.15	-3.75	4.22	-	-	-	5.06	-16.89	20.57
SD	† normal	BVN	4.34	1.85	3.54	17.82	16.43	14.91	15.14	13.23	23.63
		Frank	4.61	1.93	3.59	18.57	17.18	14.98	16.33	14.19	24.25
	§	$\text{Cln}\{0^\circ, 90^\circ\}$	4.51	1.91	3.79	19.69	18.66	15.77	13.85	17.67	24.59
		$\text{Cln}\{0^\circ, 270^\circ\}$	4.29	2.00	3.72	18.96	18.44	17.75	26.62	17.98	22.34
	beta	BVN	3.93	2.27	3.36	3.97	3.00	3.38	14.80	13.22	23.57
		Frank	4.12	2.33	3.36	4.17	3.07	3.36	16.02	14.04	24.82
		$\text{Cln}\{0^\circ, 90^\circ\}$	4.09	2.45	3.63	4.40	3.72	3.55	13.94	17.42	25.30
		$\text{Cln}\{0^\circ, 270^\circ\}$	3.90	2.41	3.61	4.23	3.18	4.24	26.48	18.94	22.44
$\sqrt{\bar{V}}$	† normal	BVN	3.96	1.75	3.36	16.77	15.49	13.55	12.37	11.02	18.24
		Frank	3.88	1.72	3.18	17.00	15.72	13.16	12.01	10.82	17.07
	§	$\text{Cln}\{0^\circ, 90^\circ\}$	3.74	1.67	2.82	16.10	14.39	10.77	10.93	8.33	13.22
		$\text{Cln}\{0^\circ, 270^\circ\}$	3.39	1.62	2.98	15.77	14.11	13.03	11.19	7.60	10.05
	beta	BVN	3.54	1.98	3.04	3.77	2.49	3.02	12.46	11.15	17.99
		Frank	3.52	1.92	2.89	3.88	2.48	2.88	12.42	10.92	17.50
		$\text{Cln}\{0^\circ, 90^\circ\}$	3.40	1.81	2.47	3.59	2.41	2.30	11.21	8.74	13.76
		$\text{Cln}\{0^\circ, 270^\circ\}$	3.13	1.80	2.70	3.51	2.24	2.84	11.68	8.08	10.21
RMSE	† normal	BVN	4.34	1.85	3.54	19.47	16.58	15.01	15.84	13.60	29.23
		Frank	4.62	1.93	3.60	19.96	17.23	15.13	17.38	14.22	28.76
	§	$\text{Cln}\{0^\circ, 90^\circ\}$	4.52	1.91	3.90	20.78	18.77	16.26	14.32	18.98	32.32
		$\text{Cln}\{0^\circ, 270^\circ\}$	4.47	2.01	3.72	21.08	18.50	17.83	27.15	24.55	28.41
	beta	BVN	4.45	4.24	5.26	-	-	-	15.34	13.97	30.29
		Frank	4.49	4.26	4.89	-	-	-	17.06	14.14	30.37
		$\text{Cln}\{0^\circ, 90^\circ\}$	4.48	4.49	5.80	-	-	-	14.19	20.13	31.71
		$\text{Cln}\{0^\circ, 270^\circ\}$	4.06	4.46	5.55	-	-	-	26.96	25.37	30.44

§: True model; †: The resulting model is the same as the TGLMM; $\text{Cln}\{\omega_1^\circ, \omega_2^\circ\}$: The $C_{13}(\cdot; \tau_{13})$ and $\{C_{12}(\cdot; \tau_{12}), C_{23|1}(\cdot; \tau_{23|1})\}$ pair copulas are Clayton rotated by ω_1 and ω_2 degrees, respectively.

Table 2: Small sample of sizes $N = 30$ simulations (10^4 replications, $n_q = 15$) from the extended trivariate vine copula mixed model with normal margins and biases, root mean square errors (RMSEs) and standard deviations (SDs), along with the square root of the average theoretical variances ($\sqrt{\bar{V}}$), scaled by 100, for the MLEs under different copula choices and margins. The true (simulated) copula distributions are the Clayton copulas rotated by 90° for both the $C_{12}(\cdot; \tau_{12})$ and $C_{13}(\cdot; \tau_{23|1})$ copulas and the Clayton copula for the $C_{13}(\cdot; \tau_{13})$ copula. The missing probability of diseased group is larger than non-diseased group, i.e., $l^{-1}(x_4) = 0.2 > l^{-1}(x_5) = 0.1$. The SDs and $\sqrt{\bar{V}}$'s for the MLEs of γ_j 's instead of σ_j 's are provided for beta margins.

	margin	copula	π_1	π_2	π_3	σ_1	σ_2	σ_3	τ_{12}	τ_{13}	$\tau_{23 1}$
Bias	† normal	BVN	0.23	-0.12	0.10	-8.23	-2.37	-1.92	-4.97	-3.36	17.51
		Frank	0.27	-0.16	-0.25	-7.81	-1.46	-2.34	-6.30	-1.07	15.89
	§	$\text{Cln}\{0^\circ, 90^\circ\}$	0.38	-0.05	0.88	-7.15	1.97	-3.90	-3.75	-6.58	20.99
		$\text{Cln}\{0^\circ, 270^\circ\}$	1.33	-0.25	0.08	-9.77	-1.73	1.39	4.01	-16.30	17.06
	beta	BVN	-1.95	-3.60	4.07	-	-	-	-4.30	-4.84	19.04
		Frank	-1.66	-3.58	3.60	-	-	-	-6.18	-1.87	18.03
		$\text{Cln}\{0^\circ, 90^\circ\}$	-1.68	-3.75	4.50	-	-	-	-2.71	-9.87	18.86
		$\text{Cln}\{0^\circ, 270^\circ\}$	-1.01	-3.78	4.24	-	-	-	3.90	-16.65	19.97
SD	† normal	BVN	4.40	1.85	3.56	18.37	16.30	14.91	15.49	13.48	23.90
		Frank	4.64	1.92	3.59	19.06	16.96	14.99	16.81	14.49	25.15
	§	$\text{Cln}\{0^\circ, 90^\circ\}$	4.61	1.92	3.81	20.28	18.54	15.72	14.08	18.15	24.63
		$\text{Cln}\{0^\circ, 270^\circ\}$	4.30	2.00	3.74	19.55	18.17	17.63	27.34	18.38	22.75
	beta	BVN	3.97	2.27	3.36	4.07	2.99	3.38	15.25	13.49	23.76
		Frank	4.13	2.33	3.37	4.25	3.04	3.36	16.69	14.45	25.94
		$\text{Cln}\{0^\circ, 90^\circ\}$	4.16	2.45	3.64	4.44	3.70	3.54	14.29	17.92	25.75
		$\text{Cln}\{0^\circ, 270^\circ\}$	3.92	2.43	3.61	4.34	3.17	4.18	27.32	19.34	22.54
$\sqrt{\bar{V}}$	† normal	BVN	4.03	1.75	3.37	17.28	15.51	13.54	12.79	11.31	18.65
		Frank	3.94	1.72	3.19	17.47	15.70	13.15	12.29	11.09	17.52
	§	$\text{Cln}\{0^\circ, 90^\circ\}$	3.80	1.67	2.83	16.57	14.39	10.81	11.09	8.44	13.48
		$\text{Cln}\{0^\circ, 270^\circ\}$	3.46	1.63	3.00	16.30	14.13	12.99	11.20	7.76	10.23
	beta	BVN	3.58	1.98	3.05	3.88	2.49	3.01	12.92	11.44	18.54
		Frank	3.56	1.92	2.89	3.99	2.47	2.87	12.74	11.23	17.82
		$\text{Cln}\{0^\circ, 90^\circ\}$	3.45	1.81	2.47	3.66	2.39	2.29	11.44	8.84	13.93
		$\text{Cln}\{0^\circ, 270^\circ\}$	3.18	1.80	2.71	3.62	2.23	2.84	11.77	8.23	10.27
RMSE	† normal	BVN	4.40	1.86	3.56	20.13	16.47	15.04	16.27	13.90	29.63
		Frank	4.65	1.93	3.60	20.60	17.02	15.17	17.95	14.53	29.75
	§	$\text{Cln}\{0^\circ, 90^\circ\}$	4.62	1.92	3.92	21.51	18.65	16.20	14.57	19.31	32.36
		$\text{Cln}\{0^\circ, 270^\circ\}$	4.50	2.01	3.74	21.86	18.25	17.68	27.63	24.57	28.44
	beta	BVN	4.43	4.25	5.28	-	-	-	15.84	14.33	30.45
		Frank	4.45	4.27	4.93	-	-	-	17.80	14.58	31.59
		$\text{Cln}\{0^\circ, 90^\circ\}$	4.49	4.48	5.79	-	-	-	14.54	20.46	31.92
		$\text{Cln}\{0^\circ, 270^\circ\}$	4.05	4.49	5.57	-	-	-	27.59	25.52	30.11

§: True model; †: The resulting model is the same as the TGLMM; $\text{Cln}\{\omega_1^\circ, \omega_2^\circ\}$: The $C_{13}(\cdot; \tau_{13})$ and $\{C_{12}(\cdot; \tau_{12}), C_{23|1}(\cdot; \tau_{23|1})\}$ pair copulas are Clayton rotated by ω_1 and ω_2 degrees, respectively.

Table 3: Small sample of sizes $N = 30$ simulations (10^4 replications, $n_q = 15$) from the extended trivariate vine copula mixed model with beta margins and biases, root mean square errors (RMSEs) and standard deviations (SDs), along with the square root of the average theoretical variances ($\sqrt{\bar{V}}$), scaled by 100, for the MLEs under different copula choices and margins. The true (simulated) copula distributions are the Clayton copulas rotated by 90° for both the $C_{12}(\cdot; \tau_{12})$ and $C_{13}(\cdot; \tau_{23|1})$ copulas and the Clayton copula for the $C_{13}(\cdot; \tau_{13})$ copula. The missing probabilities for diseased and non-diseased subjects are the same, i.e., $l^{-1}(x_4) = l^{-1}(x_5) = 0.1$. The SDs and $\sqrt{\bar{V}}$'s for the MLEs of σ_j 's instead of γ_j 's are provided for normal margins.

	margin	copula	π_1	π_2	π_3	γ_1	γ_2	γ_3	τ_{12}	τ_{13}	$\tau_{23 1}$
Bias	† normal	BVN	2.30	3.59	-2.78	-	-	-	-5.71	-2.61	14.54
		Frank	2.17	3.51	-2.89	-	-	-	-6.96	-0.70	12.33
		Cln $\{0^\circ, 90^\circ\}$	2.22	3.67	-2.52	-	-	-	-4.04	-2.62	15.72
		Cln $\{0^\circ, 270^\circ\}$	2.98	3.51	-2.79	-	-	-	1.31	-12.10	10.89
	beta	BVN	0.65	-0.02	0.10	-1.50	-0.42	-0.57	-6.48	-3.65	18.12
		Frank	0.73	-0.12	-0.12	-1.52	-0.23	-0.69	-7.85	-0.99	17.17
	§	Cln $\{0^\circ, 90^\circ\}$	0.72	-0.02	0.21	-1.54	-0.09	-1.04	-3.92	-5.90	15.13
		Cln $\{0^\circ, 270^\circ\}$	1.35	-0.24	0.21	-1.95	-0.26	-0.13	-0.33	-12.28	15.11
SD	† normal	BVN	3.50	1.59	2.77	16.26	20.91	13.08	17.04	15.65	30.27
		Frank	3.67	1.65	2.81	16.71	22.07	13.12	18.44	16.95	31.64
		Cln $\{0^\circ, 90^\circ\}$	3.68	1.64	2.91	17.45	21.86	13.91	14.80	20.88	28.19
		Cln $\{0^\circ, 270^\circ\}$	3.44	1.71	2.86	17.13	23.36	15.00	28.26	20.83	29.38
	beta	BVN	3.28	1.87	2.61	3.16	2.91	2.42	17.13	15.64	31.02
		Frank	3.40	1.97	2.62	3.25	3.08	2.39	18.87	16.91	32.42
	§	Cln $\{0^\circ, 90^\circ\}$	3.41	1.95	2.74	3.32	3.19	2.46	14.97	20.90	28.71
		Cln $\{0^\circ, 270^\circ\}$	3.21	2.02	2.70	3.27	3.20	2.85	28.91	22.01	29.84
$\sqrt{\bar{V}}$	† normal	BVN	3.25	1.42	2.63	14.95	20.02	11.74	14.03	13.32	23.34
		Frank	3.20	1.42	2.53	14.90	20.51	11.56	12.72	13.04	20.91
		Cln $\{0^\circ, 90^\circ\}$	3.13	1.34	2.34	14.48	18.15	10.00	11.35	10.33	15.90
		Cln $\{0^\circ, 270^\circ\}$	2.87	1.35	2.48	14.04	18.83	11.58	10.43	10.07	13.53
	beta	BVN	3.05	1.78	2.48	3.01	2.77	2.30	14.45	13.43	24.36
		Frank	3.03	1.77	2.39	3.01	2.86	2.23	13.67	13.25	22.08
	§	Cln $\{0^\circ, 90^\circ\}$	2.98	1.60	2.15	2.87	2.52	1.87	12.01	10.94	16.58
		Cln $\{0^\circ, 270^\circ\}$	2.77	1.68	2.39	2.78	2.63	2.29	11.70	10.87	14.47
RMSE	† normal	BVN	4.19	3.93	3.92	-	-	-	17.97	15.87	33.58
		Frank	4.27	3.88	4.03	-	-	-	19.71	16.96	33.96
		Cln $\{0^\circ, 90^\circ\}$	4.30	4.02	3.85	-	-	-	15.35	21.05	32.28
		Cln $\{0^\circ, 270^\circ\}$	4.55	3.90	3.99	-	-	-	28.29	24.09	31.33
	beta	BVN	3.34	1.87	2.61	3.50	2.94	2.48	18.32	16.07	35.93
		Frank	3.48	1.97	2.62	3.59	3.09	2.49	20.44	16.94	36.69
	§	Cln $\{0^\circ, 90^\circ\}$	3.48	1.95	2.75	3.66	3.19	2.67	15.47	21.72	32.45
		Cln $\{0^\circ, 270^\circ\}$	3.48	2.03	2.71	3.81	3.21	2.85	28.91	25.21	33.45

§: True model; †: The resulting model is the same as the TGLMM; Cln $\{\omega_1^\circ, \omega_2^\circ\}$: The $C_{13}(\cdot; \tau_{13})$ and $\{C_{12}(\cdot; \tau_{12}), C_{23|1}(\cdot; \tau_{23|1})\}$ pair copulas are Clayton rotated by ω_1 and ω_2 degrees, respectively.

Table 4: Small sample of sizes $N = 30$ simulations (10^4 replications, $n_q = 15$) from the extended trivariate vine copula mixed model with beta margins and biases, root mean square errors (RMSEs) and standard deviations (SDs), along with the square root of the average theoretical variances ($\sqrt{\bar{V}}$), scaled by 100, for the MLEs under different copula choices and margins. The true (simulated) copula distributions are the Clayton copulas rotated by 90° for both the $C_{12}(\cdot; \tau_{12})$ and $C_{13}(\cdot; \tau_{23|1})$ copulas and the Clayton copula for the $C_{13}(\cdot; \tau_{13})$ copula. The missing probability of diseased group is larger than non-diseased group, i.e., $l^{-1}(x_4) = 0.2 > l^{-1}(x_5) = 0.1$. The SDs and $\sqrt{\bar{V}}$'s for the MLEs of σ_j 's instead of γ_j 's are provided for normal margins.

	margin	copula	π_1	π_2	π_3	γ_1	γ_2	γ_3	τ_{12}	τ_{13}	$\tau_{23 1}$
Bias	† normal	BVN	2.31	3.61	-2.79	-	-	-	-6.23	-2.45	14.65
		Frank	2.20	3.54	-2.90	-	-	-	-7.42	-0.79	12.08
		Cln $\{0^\circ, 90^\circ\}$	2.22	3.70	-2.55	-	-	-	-4.31	-2.08	15.57
		Cln $\{0^\circ, 270^\circ\}$	2.99	3.54	-2.79	-	-	-	0.35	-11.53	10.19
	beta	BVN	0.72	-0.01	0.10	-1.54	-0.40	-0.62	-6.84	-3.68	17.82
		Frank	0.81	-0.11	-0.11	-1.54	-0.22	-0.73	-8.34	-1.09	16.99
	§	Cln $\{0^\circ, 90^\circ\}$	0.79	0.00	0.20	-1.60	-0.08	-1.05	-3.90	-5.64	15.04
		Cln $\{0^\circ, 270^\circ\}$	1.41	-0.23	0.23	-1.99	-0.23	-0.23	-0.77	-11.75	14.12
SD	† normal	BVN	3.55	1.58	2.78	16.73	20.95	13.20	17.76	16.29	31.53
		Frank	3.70	1.65	2.80	17.14	22.15	13.31	19.18	17.27	31.96
		Cln $\{0^\circ, 90^\circ\}$	3.74	1.64	2.92	17.87	22.03	14.09	15.23	21.30	28.90
		Cln $\{0^\circ, 270^\circ\}$	3.48	1.70	2.86	17.51	23.51	14.93	28.85	21.60	29.88
	beta	BVN	3.31	1.86	2.61	3.24	2.91	2.43	18.09	16.45	32.47
		Frank	3.43	1.97	2.62	3.33	3.08	2.40	19.93	17.41	33.82
	§	Cln $\{0^\circ, 90^\circ\}$	3.45	1.95	2.75	3.38	3.20	2.47	15.43	21.32	29.24
		Cln $\{0^\circ, 270^\circ\}$	3.24	2.02	2.72	3.35	3.25	2.80	29.75	22.91	29.87
$\sqrt{\bar{V}}$	† normal	BVN	3.31	1.41	2.62	15.49	20.03	11.73	14.52	13.76	24.73
		Frank	3.26	1.41	2.53	15.41	20.53	11.56	12.93	13.30	21.08
		Cln $\{0^\circ, 90^\circ\}$	3.21	1.34	2.35	15.03	18.28	10.06	11.59	10.53	16.16
		Cln $\{0^\circ, 270^\circ\}$	2.91	1.34	2.47	14.57	18.87	11.53	10.40	10.23	13.70
	beta	BVN	3.11	1.78	2.48	3.11	2.77	2.29	15.05	13.90	25.80
		Frank	3.09	1.77	2.39	3.11	2.86	2.22	14.04	13.54	22.91
	§	Cln $\{0^\circ, 90^\circ\}$	3.04	1.61	2.16	2.96	2.54	1.88	12.29	11.23	17.11
		Cln $\{0^\circ, 270^\circ\}$	2.83	1.68	2.39	2.88	2.66	2.27	11.91	11.19	14.98
RMSE	† normal	BVN	4.23	3.94	3.93	-	-	-	18.83	16.48	34.76
		Frank	4.30	3.90	4.03	-	-	-	20.56	17.29	34.17
		Cln $\{0^\circ, 90^\circ\}$	4.35	4.05	3.88	-	-	-	15.83	21.40	32.83
		Cln $\{0^\circ, 270^\circ\}$	4.58	3.92	3.99	-	-	-	28.85	24.48	31.57
	beta	BVN	3.38	1.86	2.62	3.59	2.94	2.50	19.34	16.85	37.04
		Frank	3.52	1.97	2.63	3.67	3.09	2.51	21.60	17.45	37.85
	§	Cln $\{0^\circ, 90^\circ\}$	3.54	1.95	2.76	3.75	3.20	2.68	15.91	22.05	32.88
		Cln $\{0^\circ, 270^\circ\}$	3.53	2.03	2.73	3.90	3.26	2.81	29.76	25.74	33.04

§: True model; †: The resulting model is the same as the TGLMM; Cln $\{\omega_1^\circ, \omega_2^\circ\}$: The $C_{13}(\cdot; \tau_{13})$ and $\{C_{12}(\cdot; \tau_{12}), C_{23|1}(\cdot; \tau_{23|1})\}$ pair copulas are Clayton rotated by ω_1 and ω_2 degrees, respectively.