*Biometrics*, 2025, 81(2), ujaf037
https://doi.org/10.1093/biomtc/ujaf037
**Biometric Methodology**

OXFORD

# Vine copula mixed models for meta-analysis of diagnostic accuracy studies without a gold standard

Aristidis K. Nikoloulopoulos *

School of Engineering, Mathematics and Physics, University of East Anglia, Norwich NR47TJ, United Kingdom

*Corresponding author: Aristidis K. Nikoloulopoulos, School of Engineering, Mathematics and Physics, University of East Anglia, Norwich NR47TJ, United Kingdom (a.nikoloulopoulos@uea.ac.uk).



## ABSTRACT

Numerous statistical models have been proposed for conducting meta-analysis of diagnostic accuracy studies when a gold standard is available. However, in real-world scenarios, the gold standard test may not be perfect due to several factors such as measurement error, non-availability, invasiveness, or high cost. A generalized linear mixed model (GLMM) is currently recommended to account for an imperfect reference test. We propose vine copula mixed models for meta-analysis of diagnostic test accuracy studies with an imperfect reference standard. Our general models include the GLMM as a special case, can have arbitrary univariate distributions for the random effects, and can provide tail dependencies and asymmetries. Our general methodology is demonstrated with an extensive simulation study and illustrated by insightfully re-analyzing the data of a meta-analysis of the Papanicolaou test that diagnoses cervical neoplasia. Our study suggests that there can be an improvement on GLMM and makes the argument for moving to vine copula random effects models.

**KEYWORDS:** imperfect reference test; meta-analysis; mixed models; vine copulas.

## 1 INTRODUCTION

The rise of evidence-based medicine has resulted in increased focus on meta-analytic studies of diagnostic test accuracy. A meta-analysis of diagnostic test accuracy studies combines information from different studies. It provides an integrated analysis with more statistical power to detect an accurate diagnostic test than an analysis based on a single study (Normand, 1999). As diagnostic test, accuracy is commonly measured by a pair of indices such as sensitivity (the probability that an actual positive will test positive) and specificity (the probability that an actual negative will test negative), synthesis of diagnostic test accuracy studies is the most common medical application of multivariate meta-analysis (Jackson et al., 2011). Most existing meta-analysis models and methods have mainly focused on a single test when a perfect reference standard is available (eg, Rutter and Gatsonis, 2001; Reitsma et al., 2005; Chu and Cole, 2006; Nikoloulopoulos, 2015).

Nevertheless, in practice, the reference test may be imperfect due to measurement error, non-existence, invasiveness, or the high cost of a gold standard. This is the case in the meta-analysis of 59 studies to evaluate the accuracy of the Papanicolaou (Pap) test that diagnoses cervical neoplasia (Fahey et al., 1995; Liu et al., 2015). As acknowledged in Liu et al. (2015), the literature on meta-analytic studies of diagnostic test accuracy, when a gold standard is unavailable, is very limited. Chu et al. (2009) proposed a 5-variate generalized linear mixed model (GLMM) to account for heterogeneity in test accuracies across studies by treating the disease prevalence (the probability of those with the disease), sensitivities, and specificities of the index and reference tests as random effects. Their model assumes independent multinomial distributions for the counts of each combination of the index and reference test results, conditional on the 5-variate normally distributed transformed latent disease prevalence, sensitivities, and specificities of the index and reference tests in each study. Dendukuri et al. (2012) extended the hierarchical summary receiver operating characteristic (HSROC) of Rutter and Gatsonis (2001) to the situation where no gold standard test is available. Both the GLMM and HSROC have the advantage that they account for heterogeneity across studies and allow for dependence between the index and reference test. As shown by Liu et al. (2015), who re-analyzed the existing meta-data on the diagnosis of cervical neoplasia, the GLMM and HSROC are closely related, and some of their submodels with fewer random effects are equivalent.

Nevertheless, the 5-variate normal distribution of the transformed latent proportions in the GLMM has restricted properties, that is, a linear correlation structure and normal margins that might lead to biased meta-analytic estimates of diagnostic test accuracy. In order to create a flexible distribution to model the random effects, we exploit the use of regular vine copulas (Bedford and Cooke, 2002) as other parametric copulas such as Archimedean, nested Archimedean, and elliptical copulas have limited dependence (Nikoloulopoulos, 2013). Regular vine copulas are suitable for high-dimensional data (Erhardt et al., 2015); hence, given the low dimension, we use their boundary case, namely a D-vine copula. D-vine copulas have be-

Received: July 30, 2024; Revised: January 10, 2025; Accepted: March 12, 2025

come important in many application areas, such as finance (Aas et al., 2009; Nikoloulopoulos et al., 2012) and biological sciences (Killiches and Czado, 2018; Barthel et al., 2019; Hoque et al., 2022; Sahin and Czado, 2024), to name just a few, in order to deal with dependence in the joint tails. Another boundary case of regular vine copulas is the canonical vine copula, but this parametric family of copulas is suitable if there exists a pilot variable that drives the dependence among the variables (Nikoloulopoulos and Joe, 2015; Erhardt and Czado, 2018; Kadhem and Nikoloulopoulos, 2021), which apparently is not the case in this application area.

We propose a copula mixed model as an extension of the GLMM by using a D-vine copula representation of the random effects distribution with both normal and beta margins. We assume independent multinomial distributions for the counts of each combination of the index and reference test results, conditional on the latent disease prevalence, sensitivities, and specificities of the index and reference tests in each study. The proposed model (1) includes the 5-variate GLMM (Chu et al., 2009) as a special case, (2) can have arbitrary univariate distributions for the random effects, and (3) can provide tail dependencies and asymmetries.

The remainder of the paper proceeds as follows. Section 2 summarizes the standard 5-variate GLMM for synthesis of diagnostic test accuracy studies without a gold standard. Section 3 introduces the 5-variate D-vine copula mixed model for meta-analysis of diagnostic studies without a gold standard, and provides computational details for maximum likelihood (ML) estimation. Section 4 insightfully re-analyzes the data from the meta-analysis of the Pap test that diagnoses cervical neoplasia. Section 5 studies the reliability of our methodology to select the necessary latent variables (random effects) and gauge the small-sample efficiency and robustness of the ML estimation of the proposed D-vine copula mixed model. We conclude with some discussion in Section 6.

## 2 THE 5-VARIATE GLMM

In this section, we summarize the 5-variate GLMM in Chu et al. (2009). Before that, we first introduce the notation we use. The data are $y_{ijk},\ i = 1, \ldots, N,\ j = 0, 1,\ k = 0, 1$, where $i$ is an index for the individual studies, $j$ is an index for the index test outcome (0: negative; 1: positive), and $k$ is an index for the imperfect reference test outcome (0: negative; 1: positive). Each cell in Table 1 provides the cell frequency corresponding to a combination of the index and reference test outcomes in study $i$.

The within-study model assumes that the counts $(Y_{i11}, Y_{i10}, Y_{i01}, Y_{i00})$ of each combination of test results are multinomially distributed given the transformed latent disease prevalence $X_1$, the transformed latent sensitivities $X_2$ and $X_3$ for the index and reference test, and the transformed latent specificities $X_4$ and $X_5$ for the index and reference test, respectively, viz.,

$$\begin{aligned}&(Y_{11i}, Y_{10i}, Y_{01i}, Y_{00i})|(X_1, X_2, X_3, X_4, X_5) = (x_1, x_2, x_3, x_4, x_5)\\ &\sim \mathcal{M}_4(y_{i++}, p_{11}, p_{10}, p_{01}, p_{00}),\end{aligned} \tag{1}$$

where $p_{jk}$ are the cell latent probabilities in Table 1, and $\mathcal{M}_T(n, p_1, \ldots, p_d)$ is shorthand notation for the multinomial distribution; $d$ is the number of cells, $n$ is the number of observations, and $(p_1, \ldots, p_d)$ with $p_1 + \ldots + p_d = 1$ is the $d$-dimensional vector of success probabilities. Under the assumption that the 2 tests are independent conditional on the true disease status, the cell latent probabilities are the following functions of the disease prevalence, sensitivities, and specificities for the index and reference test:

$$\begin{aligned}p_{11} &= l^{-1}(x_1)l^{-1}(x_2)l^{-1}(x_3) +\\ &\quad \{1 - l^{-1}(x_1)\}\{1 - l^{-1}(x_4)\}\{1 - l^{-1}(x_5)\}\\ p_{10} &= l^{-1}(x_1)l^{-1}(x_2)\{1 - l^{-1}(x_3)\} +\\ &\quad \{1 - l^{-1}(x_1)\}\{1 - l^{-1}(x_4)\}l^{-1}(x_5)\\ p_{01} &= l^{-1}(x_1)\{1 - l^{-1}x_2)\}l^{-1}(x_3) +\\ &\quad \{1 - l^{-1}(x_1)\}l^{-1}(x_4)\{1 - l^{-1}(x_5)\},\end{aligned} \tag{2}$$

where $l(p)$ is a link function as the commonly used $\text{logit}(p) = \log(\frac{p}{1-p})$. For further explanation on the structure of equations in (2), we refer the interested reader to Gart and Buck (1966).

The between-studies model assumes that $\mathbf{X}$ follows a multivariate normal (MVN) distribution with mean vector $\boldsymbol{\mu} = (l(\pi_1), l(\pi_2), l(\pi_3), l(\pi_4), l(\pi_5))$ and variance-covariance matrix:

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 & \rho_{14}\sigma_1\sigma_4 & \rho_{15}\sigma_1\sigma_5\\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 & \rho_{24}\sigma_2\sigma_4 & \rho_{25}\sigma_2\sigma_5\\ \rho_{13}\sigma_1\sigma_3 & \rho_{23}\sigma_2\sigma_3 & \sigma_3^2 & \rho_{34}\sigma_3\sigma_4 & \rho_{35}\sigma_3\sigma_5\\ \rho_{14}\sigma_1\sigma_4 & \rho_{24}\sigma_2\sigma_4 & \rho_{34}\sigma_3\sigma_4 & \sigma_4^2 & \rho_{45}\sigma_4\sigma_5\\ \rho_{15}\sigma_1\sigma_5 & \rho_{25}\sigma_2\sigma_5 & \rho_{35}\sigma_3\sigma_5 & \rho_{45}\sigma_4\sigma_5 & \sigma_5^2 \end{pmatrix}.$$

Here, $\pi_1$ is the meta-analytic parameter for the prevalence, $\pi_2$ and $\pi_3$ are the meta-analytic parameters for the sensitivity of the index and the reference test, respectively, and $\pi_4$ and $\pi_5$ are the meta-analytic parameters for the specificity of the index and the reference test, respectively. In addition, the variance parameters $\sigma_t^2,\ t = 1, \ldots, 5$ denote the between-study heterogeneity in disease prevalence, sensitivities, and specificities for the index and reference test, and the off-diagonal parameters $\rho_{t_1t_2} : 1 \leq t_1 < t_2 \leq 5$ denote the pairwise correlations among the transformed latent prevalence, sensitivities, and specificities (random effects).

## 3 THE 1-TRUNCATED D-VINE COPULA MIXED MODEL

In this section, we introduce the 1-truncated D-vine copula mixed model for the meta-analysis of diagnostic accuracy stud-

**TABLE 1** Data from an individual study in a 2 × 2 table.

| | Index test | |
|---|---|---|
| **Reference test** | **Positive** | **Negative** |
| Positive | $y_{i11}$ | $y_{i10}$ |
| Negative | $y_{i01}$ | $y_{i00}$ |

ies without a gold standard and discuss its relationship with the GLMM in the preceding section. Before that, the first subsection has some background on vine copula models. We complete this section with details on ML estimation.

### 3.1 Overview and relevant background for vine copulas

A copula is a multivariate cumulative distribution function (cdf) with uniform $U(0, 1)$ margins (Joe, 2014). If $F$ is a $\mathcal{T}$-variate cdf with univariate margins $F_1, \ldots, F_{\mathcal{T}}$, then there is a copula $C$ such that

$$F(x_1, \ldots, x_{\mathcal{T}}) = C\Big(F_1(x_1), \ldots, F_d(x_{\mathcal{T}})\Big).$$

The copula is unique if $F_1, \ldots, F_{\mathcal{T}}$ are continuous. If $F$ is continuous and $(Y_1, \ldots, Y_{\mathcal{T}}) \sim F$, then the unique copula is the distribution of $(U_1, \ldots, U_{\mathcal{T}}) = (F_1(Y_1), \ldots, F_{\mathcal{T}}(Y_{\mathcal{T}}))$ leading to

$$C(u_1, \ldots, u_T) = F\Big(F_1^{-1}(u_1), \ldots, F_{\mathcal{T}}^{-1}(u_{\mathcal{T}})\Big),$$
$$0 \le u_t \le 1, t = 1, \ldots, \mathcal{T},$$

where $F_t^{-1}$ are inverse cdfs (Nikoloulopoulos and Joe, 2015). In particular, if $\Phi_{\mathcal{T}}(\cdot; \mathbf{R})$ is the MVN cdf with correlation matrix

$$\mathbf{R} = (\rho_{t_1 t_2} : 1 \le t_1 < t_2 \le \mathcal{T})$$

and $N(0,1)$ margins, and $\Phi$ is the univariate standard normal cdf, then the MVN copula is

$$C(u_1, \ldots, u_{\mathcal{T}}) = \Phi_{\mathcal{T}}\Big(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_{\mathcal{T}}); \mathbf{R}\Big). \quad (3)$$

Copulas have become useful for flexible modeling of multivariate data when the variables are non-normal. In particular, there are copula families that can lead to more dependence in the upper or lower joint tail than with the MVN copula and these are important for dependence among extreme values.

The $\mathcal{T}$-dimensional D-vine copulas can cover flexible dependence structures, different from assuming simple linear correlation structures, tail independence and normality (Joe et al., 2010). They are built via successive mixing from $\mathcal{T}(\mathcal{T}-1)/2$ bivariate linking copulas on levels. For the $\mathcal{T}$-dimensional D-vine, the bivariate pairs at level 1 are $X_t, X_{t+1}$, for $t = 1, \ldots, \mathcal{T}-1$, and for level $\ell$ $(2 \le \ell < \mathcal{T})$, the (conditional) bivariate pairs are $X_t, X_{t+\ell}|X_{t+1}, \ldots, X_{t+\ell-1}$ for $t = 1, \ldots, \mathcal{T}-\ell$. That is, for the D-vine, conditional bivariate copulas are specified for variables $t$ and $t+\ell$ given the variables indexed in between (Nikoloulopoulos et al., 2012). In Figure 1, the 5-dimensional D-vine copula with $\mathcal{T} = 5$ variables and 4 trees/levels is depicted.

Joe et al. (2010) have shown that in order for a vine copula to have (tail) dependence for all bivariate margins, it is only necessary for the bivariate copulas in level 1 to have (tail) dependence, and it is not necessary for the conditional bivariate copulas in levels $2, \ldots, \mathcal{T}-1$ to have (tail) dependence. That provides the theoretical justification for the idea to model the dependence in the first level and then just use the independence copulas to model conditional dependence at higher levels without sacrificing the tail dependence of the vine copula distribution. In line with our previous contributions in copula mixed models (Nikoloulopoulos, 2015, 2017, 2018a,b, 2019, 2020a,b, 2022, 2024b), we use bivariate parametric copulas with different tail dependence behavior, namely the bivariate normal (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°). Note in passing that vine copulas have as a special case the MVN copula in (3), if all the bivariate parametric copulas are BVN (Aas et al., 2009).

### 3.2 The 1-truncated D-vine copula mixed model with normal and beta margins

Here, we generalize the GLMM in Chu et al. (2009) by proposing a model that links the random effects using a 1-truncated D-vine copula rather than the MVN distribution. The within-study model is the same as in the standard GLMM; see (1).

For the between-studies model, there are different latent variables $(X_1, X_2, X_3, X_4, X_5)$, but they are dependent. To model the dependence among the latent variables $X_t$, $t = 1, \ldots, 5$, we employ copulas. The power of copulas for dependence modeling is due to the dependence structure being considered separate from the univariate margins. For $t = 1, \ldots, 5$ denote the univariate cdf of $X_t$ by $F\big(\cdot; l(\pi_t), \delta_t\big)$, whereas in Section 2, $\pi_t$, $t = 1, \ldots, 5$ are the meta-analytic parameters of the proportions of

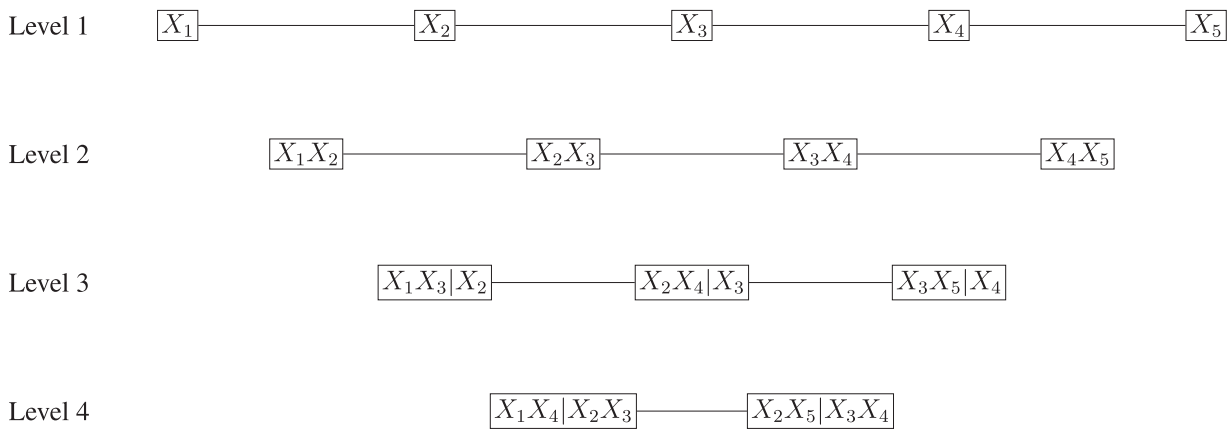


**FIGURE 1** Graphical representation of the D-vine copula with 5 variables and 4 levels.

**TABLE 2** The choices of the $F(\cdot; l(\pi), \delta)$ and $l$ in the copula mixed model.

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

interest, but now the univariate parameters $\delta_t$, $t = 1, \ldots, 5$ denote the between-study heterogeneity in disease prevalence, sensitivities, and specificities of the index and reference test. The choices of the $F(\cdot; l(\pi), \delta)$ and $l$ are given in Table 2. If the $\text{Beta}(\pi, \gamma)$ distribution is used for the marginal modeling of the latent proportions, then one does not have to transform the prevalence, sensitivities, and specificities and can work on the original scale.

In multivariate models with copulas, a copula or multivariate uniform distribution is combined with a set of univariate margins. That is, if a 5-dimensional parametric family of copulas $C_5(\cdot; \boldsymbol{\theta})$ is combined with the parametric model $F(\cdot; l(\pi_t), \delta_t)$, then

$$C_5\Big(F(x_1; l(\pi_1), \delta_1), F(x_2; l(\pi_2), \delta_2), F(x_3; l(\pi_3), \delta_3), F(x_4; l(\pi_4), \delta_4), F(x_5; l(\pi_5), \delta_5); \boldsymbol{\theta}\Big)$$

is a multivariate parametric model with univariate margins $F(\cdot; l(\pi_t), \delta_t)$, $t = 1, \ldots, 5$. This is equivalent to assuming that the latent variables $X_t$, $t = 1, \ldots, 5$ have been transformed to standard uniform latent variables $U_t = F(X_t; l(\pi_t), \delta_t)$, $t = 1, \ldots, 5$. So we assume that $(U_1, U_2, U_3, U_4, U_5)$ is a 5-dimensional random vector, where $U_t \sim U(0, 1)$. The joint cdf is then given by $C_5(u_1, u_2, u_3, u_4, u_5; \boldsymbol{\theta})$, where $C_5(\cdot; \boldsymbol{\theta})$ is a 5-dimensional 1-truncated D-vine copula, with copula parameter vector $\boldsymbol{\theta} = (\theta_{12}, \theta_{23}, \theta_{34}, \theta_{45})$. The 1-truncated D-vine copula has 4 parametric bivariate copulas $C(\cdot; \theta_{12}), C(\cdot; \theta_{23}), C(\cdot; \theta_{34})$, and $C(\cdot; \theta_{45})$ that link $X_1$ with $X_2$, $X_2$ with $X_3$, $X_3$ with $X_4$, and $X_4$ with $X_5$, respectively, in the first level of the vine and independence copulas in all the remaining levels of the vine (truncated after the first level). If one is restricted to the first level, then the result is a Markov tree dependence structure where 2 variables not connected by an edge are conditionally independent given the variables between them. Figure 2 depicts the graphical representation of the 1-truncated D-vine copula model. This truncation, as per the terminology in Brechmann et al. (2012), offers a substantial reduction of the dependence parameters. In our case, there are 6 (conditional) dependence parameters less, which is extremely useful for estimation purposes given the typically small number of primary studies involved in meta-analysis. To this end, the stochastic representation of the between-studies model takes the form

$$\Big(F(X_1; l(\pi_1), \delta_1), \ldots, F(X_5; l(\pi_5), \delta_5)\Big) \sim C_5(\cdot; \boldsymbol{\theta}). \quad (4)$$

Let $c(u, v; \theta) = \frac{\partial C(u,v;\theta)}{\partial u \partial v}$ be a bivariate copula density. Then, the 5-dimensional 1-truncated D-vine copula density is decomposed in a simple manner by multiplying the bivariate copulas densities in the nodes of the tree in Figure 2, as indicated below

$$c_5\Big(u_1, u_2, u_3, u_4, u_5; \boldsymbol{\theta}\Big) = c\Big(u_1, u_2; \theta_{12}\Big)c\Big(u_2, u_3; \theta_{23}\Big) c\Big(u_3, u_4; \theta_{34}\Big)c\Big(u_4, u_5; \theta_{45}\Big). \quad (5)$$

Note that for a 5-dimensional D-vine copula density, there are $\frac{5!}{2}$ distinct permutations of the variables (Aas et al., 2009). To be concrete in the exposition of the theory, we use the permutation in Figure 2; the theory though also applies to the other permutations. The models in (1) and (4) together specify a 1-truncated D-vine copula mixed model with joint likelihood

$$L(\pi_1, \pi_2, \pi_3, \pi_4, \pi_5, \delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \boldsymbol{\theta}) = \prod_{i=1}^{N} \int_{[0,1]^5} g\Big(y_{i11}, y_{i10}, y_{i01}, y_{i00}; y_{i++}, p_{11}, p_{10}, p_{01}\Big) c_5(u_1, u_2, u_3, u_4, u_5; \boldsymbol{\theta}) du_1 du_2 du_3 du_4 du_5, \quad (6)$$

where $p_{jk}$ are as in (2) with $x_t = F^{-1}(u_t; l(\pi_t), \delta_t)$, $t = 1, \ldots, 5$.

Our general statistical model allows for the selection of copulas and margins independently, that is, there are no constraints in the choices of parametric copulas and margins. Note in passing that when the univariate distribution of the random effects is the $N(\mu, \sigma)$ distribution and all the bivariate copulas are BVN with copula (correlation) parameters $\rho_{12}, \rho_{23}, \rho_{34}, \rho_{45}$, the resulting random effects distribution is the MVN with mean vector $\boldsymbol{\mu}$ and variance-covariance matrix $\boldsymbol{\Sigma}$ where $\rho_{13} = \rho_{12}\rho_{23}, \rho_{14} = \rho_{12}\rho_{24}, \rho_{15} = \rho_{14}\rho_{45}, \rho_{24} = \rho_{23}\rho_{34}, \rho_{25} = \rho_{24}\rho_{45}, \rho_{35} = \rho_{34}\rho_{45}$. Hence, the 1-truncated D-vine copula mixed model has as special the GLMM with a structured correlation matrix. The above pairwise correlations are deduced using the partial correlation vine parametrization (Joe, 2014, page 119) of the MVN copula, which consists of algebraically independent pairwise correlations (first level) and partial correlations (higher levels).

### 3.3 ML estimation and computational details

Estimation of the model parameters can be approached by the standard ML method, by maximizing the logarithm of the joint likelihood. The estimated parameters can be obtained by using a quasi-Newton method applied to the logarithm of the joint likelihood. We employ a quasi-Newton method as an alternative to Newton's method in order to achieve faster computation without having to calculate the Hessian matrix every time. Therefore, the quasi-Newton minimization with an input function of the negative log-likelihood to be minimized provides the output point of minimum and the inverse Hessian at the point of minimum.

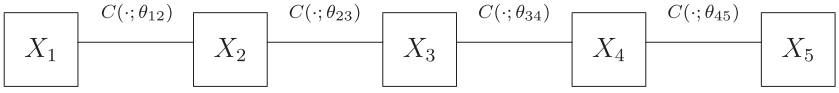


**FIGURE 2** Graphical representation of the 5-dimensional 1-truncated D-vine copula model.

For the 1-truncated D-vine copula mixed model, numerical evaluation of the joint probability mass function (pmf) can be achieved with the following steps:

(1) Calculate Gauss-Legendre quadrature points $\{u_q : q = 1, \ldots, N_q\}$ and weights $\{w_q : q = 1, \ldots, N_q\}$ in terms of standard uniform.
(2) Convert from independent uniform random variables $\{u_{q_1} : q_1 = 1, \ldots, N_q\}$, $\{u_{q_2} : q_2 = 1, \ldots, N_q\}$, $\{u_{q_3} : q_3 = 1, \ldots, N_q\}$, and $\{u_{q_4} : q_4 = 1, \ldots, N_q\}$, $\{u_{q_5} : q_5 = 1, \ldots, N_q\}$ to dependent uniform random variables that have a 1-truncated D-vine distribution $C_5(\cdot; \boldsymbol{\theta})$:
1: $v_{q_1} = u_{q_1}$
2: $v_{q_2} = C^{-1}(u_{q_2}|v_{q_1}; \theta_{12})$
3: $v_{q_3} = C^{-1}(u_{q_3}|v_{q_2}; \theta_{23})$
4: $v_{q_4} = C^{-1}(u_{q_4}|v_{q_3}; \theta_{34})$
5: $v_{q_5} = C^{-1}(u_{q_5}|v_{q_4}; \theta_{45})$,
where $C^{-1}(v|u; \theta)$ is the inverse conditional bivariate copula cdf. The method is based on the simulation algorithm of a 1-truncated D-vine copula (Joe, 2014), where as input, instead of independent uniform variates, it uses the independent quadrature points.
(3) Numerically evaluate the joint pmf

$$\int_{[0,1]^5} g\left(y_{i11}, y_{i10}, y_{i01}, y_{i00}; y_{i++}, p_{11}, p_{10}, p_{01}\right) c(u_1, u_2, u_3, u_4, u_5; \boldsymbol{\theta}) du_1 du_2 du_3 du_4 du_5$$

in a quintuple sum:

$$\sum_{q_1=1}^{N_q}\sum_{q_2=1}^{N_q}\sum_{q_3=1}^{N_q}\sum_{q_4=1}^{N_q}\sum_{q_5=1}^{N_q} w_{q_1} w_{q_2} w_{q_3} w_{q_4} w_{q_5} g\left(y_{i11}, y_{i10}, y_{i01}, y_{i00}; y_{i++}, p_{11}, p_{10}, p_{01}\right),$$

where $p_{11}, p_{10}, p_{01}$ are calculated as in (2) with $x_t = F^{-1}(v_{q_t}; \pi_t, \delta_t)$.

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.

## 4 META-ANALYSIS OF THE PAPANICOLAOU TEST

In this section, we illustrate the proposed methodology by insightfully re-analyzing the data of a meta-analysis of the Pap test that diagnoses cervical neoplasia (Fahey et al., 1995). These data are comprised of $N = 59$ studies that were published between January 1984 and March 1992. The diagnostic accuracy of the Pap test (ie, index test) is evaluated by comparing with the histology test (ie, reference test), which is not a perfect test (Fahey et al., 1995). These data have been previously analyzed by Liu et al. (2015), who have fitted the GLMM and the HSROC model. It was established that these 2 models lead to equivalent submodels with fewer random effects, and hence identical inferences. Note in passing that the GLMM is a special case of our model when all the bivariate copulas are BVN and the univariate distribution of the random effects is the $N(\mu, \sigma^2)$ distribution as discussed in Section 2.

To avoid over-fitting the data with an excess of random effects (latent variables), we use an all possible random effects procedure based on information criteria. We start with the initial assumption, that all bivariate linking copulas are BVN copulas, that is the starting model is the GLMM. We use both Akaike's information criterion (AIC) and the Bayesian information criterion (BIC) to select which random effects provide the most parsimonious fit as the number of parameters is not the same among the random effect models. The goal is to find the "best" subset of the random effects. We refer to models with 1 random effect $X_t$ as submodel A.$t$., with 2 random effects $(X_{t_1}, X_{t_2} : 1 \le t_1 < t_2 \le 5)$ as submodel B.$s$, $s = 1, \ldots, 10$, with 3 random effects $(X_{t_1}, X_{t_2}, X_{t_3} : 1 \le t_1 < t_2 < t_3 \le 5)$ as submodel C.$s$, with 4 random effects $(X_{t_1}, X_{t_2}, X_{t_3}, X_{t_4} : 1 \le t_1 < t_2 < t_3 < t_4 \le 5)$ as submodel D.$t$, and with the all 5 random effects $(X_1, X_2, X_3, X_4, X_5)$ in (1) and (4) as model E. The full specification of the submodels with fewer random effects is provided in Web Appendix A.

For the data on the Pap and histology tests, both AIC and BIC in Table 3 suggest the use of the submodel C.2, which includes random effects for the prevalence $X_1$, sensitivity $X_2$, and specificity $X_4$ of the Pap test, and fixed effects for the sensitivity $\pi_3$ and specificity $\pi_5$ of the histology test (reference test). That is, the within-study and between-studies models in (1) and (4) reduce to

$$\begin{aligned}&(Y_{11i}, Y_{10i}, Y_{01i}, Y_{00i})|(X_1, X_2, X_4) = (x_1, x_2, x_4)\\ &\sim \mathcal{M}_4(y_{i++}, p_{11}, p_{10}, p_{01}, p_{00}),\end{aligned} \tag{7}$$

where

$$\begin{aligned}p_{11} &= l^{-1}(x_1) l^{-1}(x_2)\pi_3 + \{1 - l^{-1}(x_1)\}\{1 - l^{-1}(x_4)\}(1 - \pi_5)\\ p_{10} &= l^{-1}(x_1) l^{-1}(x_2)(1 - \pi_3) + \{1 - l^{-1}(x_1)\}\{1 - l^{-1}(x_4)\}\pi_5\\ p_{01} &= l^{-1}(x_1)\{1 - l^{-1}x_2)\}\pi_3 + \{1 - l^{-1}(x_1)\} l^{-1}(x_4)(1 - \pi_5),\end{aligned} \tag{8}$$

and

$$\left(F(X_1; l(\pi_1), \delta_1), F(X_2; l(\pi_2), \delta_2), F(X_4; l(\pi_4), \delta_4)\right) \sim C_3(\cdot; \boldsymbol{\theta}), \tag{9}$$

respectively, where $C_3(\cdot; \boldsymbol{\theta})$ is a trivariate 1-truncated D-vine copula with dependence parameter vector $\boldsymbol{\theta} = (\theta_{12}, \theta_{24})$. It has 2 parametric bivariate copulas $C(\cdot; \theta_{12})$ and $C(\cdot; \theta_{24})$ that link $X_1$ with $X_2$ and $X_2$ with $X_4$, respectively, in the first level of the vine and independence copulas in all the remaining levels of the vine (truncated after the first level). Consequently, the joint likelihood of the models in (7) and (9) has the reduced form

$$\begin{aligned}&L(\pi_1, \pi_2, \pi_3, \pi_4, \pi_5, \delta_1, \delta_2, \delta_4, \boldsymbol{\theta}) =\\ &\prod_{i=1}^{N}\int_{[0,1]^3} g\left(y_{i11}, y_{i10}, y_{i01}, y_{i00}; y_{i++}, p_{11}, p_{10}, p_{01}\right)\\ &c(u_1, u_2; \theta_{12}) c(u_2, u_4; \theta_{24}) du_1 du_2 du_4,\end{aligned} \tag{10}$$

where $p_{jk}$ are as in (8) with $x_t = F^{-1}(u_t; l(\pi_t), \delta_t)$, $t = 1, 2, 4$.

After determining the random effects, we incorporate both normal and beta margins along with copulas with lower or up-

**TABLE 3** Akaike's information criterion (AIC) and the Bayesian information criterion (BIC) from all the fitted random effects models. We refer to models with 1 random effect $X_t$ as submodel A.$t$., with 2 random effects $(X_{t_1}, X_{t_2} : 1 \le t_1 < t_2 \le 5)$ as submodel B.$s$, $s = 1, \ldots, 10$, with 3 random effects $(X_{t_1}, X_{t_2}, X_{t_3} : 1 \le t_1 < t_2 < t_3 \le 5)$ as submodel C.$s$, with 4 random effects $(X_{t_1}, X_{t_2}, X_{t_3}, X_{t_4} : 1 \le t_1 < t_2 < t_3 < t_4 \le 5)$ as submodel D.$t$, and with the all 5 random effects $(X_1, X_2, X_3, X_4, X_5)$ in (1) and (4) as model E. The selected model, that is, the model with the smallest information criteria, is boldfaced.

| Random effects models | $-\ell$ | AIC | BIC |
|---|---|---|---|
| A.1 $(X_1)$ | 9792.93 | 19597.86 | 19596.49 |
| A.2 $(X_2)$ | 10199.53 | 20411.06 | 20409.69 |
| A.3 $(X_3)$ | 10388.13 | 20788.27 | 20786.89 |
| A.4 $(X_4)$ | 10295.68 | 20603.37 | 20601.99 |
| A.5 $(X_5)$ | 10213.40 | 20438.79 | 20437.42 |
| B.1 $(X_1, X_2)$ | 9259.70 | 18535.40 | 18533.57 |
| B.2 $(X_1, X_3)$ | 9539.64 | 19095.27 | 19093.44 |
| B.3 $(X_1, X_4)$ | 9539.34 | 19094.69 | 19092.85 |
| B.4 $(X_1, X_5)$ | 9259.19 | 18534.38 | 18532.55 |
| B.5 $(X_2, X_3)$ | 9430.84 | 18877.68 | 18875.85 |
| B.6 $(X_2, X_4)$ | 10037.44 | 20090.87 | 20089.04 |
| B.7 $(X_2, X_5)$ | 9722.67 | 19461.34 | 19459.50 |
| B.8 $(X_3, X_4)$ | 10072.07 | 20160.13 | 20158.30 |
| B.9 $(X_3, X_5)$ | 10081.92 | 20179.84 | 20178.00 |
| B.10 $(X_4, X_5)$ | 9332.84 | 18681.68 | 18679.85 |
| C.1 $(X_1, X_2, X_3)$ | 9066.02 | 18152.05 | 18149.76 |
| **C.2 $(\mathbf{X_1}, \mathbf{X_2}, \mathbf{X_4})$** | 9056.26 | **18132.53** | **18130.24** |
| C.3 $(X_1, X_2, X_5)$ | 9254.62 | 18529.24 | 18526.95 |
| C.4 $(X_1, X_3, X_4)$ | 9538.83 | 19097.66 | 19095.37 |
| C.5 $(X_1, X_3, X_5)$ | 9060.31 | 18140.62 | 18138.33 |
| C.6 $(X_1, X_4, X_5)$ | 9062.52 | 18145.03 | 18142.74 |
| C. 7 $(X_2, X_3, X_4)$ | 9200.49 | 18420.98 | 18418.69 |
| C.8 $(X_2, X_3, X_5)$ | 9340.18 | 18700.36 | 18698.07 |
| C.9 $(X_2, X_4, X_5)$ | 9198.71 | 18417.41 | 18415.12 |
| C.10 $(X_3, X_4, X_5)$ | 9258.81 | 18537.62 | 18535.33 |
| D.1 $(X_1, X_2, X_3, X_4)$ | 9056.94 | 18137.87 | 18135.12 |
| D.2 $(X_1, X_2, X_3, X_5)$ | 9090.45 | 18204.91 | 18202.16 |
| D.3 $(X_1, X_2, X_4, X_5)$ | 9055.70 | 18135.40 | 18132.65 |
| D.4 $(X_1, X_3, X_4, X_5)$ | 9064.27 | 18152.54 | 18149.79 |
| D.5 $(X_2, X_3, X_4, X_5)$ | 9076.88 | 18177.76 | 18175.01 |
| E $(X_1, X_2, X_3, X_4, X_5)$ | 9053.22 | 18134.44 | 18131.23 |

$X_1$: random effect for the prevalence; $X_2$: random effect for the sensitivity of the Pap test; $X_3$: random effect for the sensitivity of the histology test; $X_4$: random effect for the specificity of the Pap test; $X_5$: random effect for the specificity of the histology test.

per tail dependence necessary to account for more probability in one or both joint tails. We fit the trivariate 1-truncated D-vine copula mixed model for all different pair copulas and univariate marginal distributions. We use the decomposition of the vine copula density in (5), as different decompositions lead to similar results as long as the "best" permutation for D-vines consists of choosing and connecting the most dependent pairs (Nikoloulopoulos et al., 2012). This is the pair $(X_2, X_4)$ of the Pap test. In our general statistical model, there are no constraints in the choices of the parametric marginal or pair-copula distributions. This is one of the limitations of the GLMM where all the pair copulas are BVN and marginal distributions are normal. However, for ease of interpretation, we do not mix pair-copulas or margins. To make it easier to compare strengths of dependence, we convert the BVN, Frank, and rotated Clayton estimated copula parameters to Kendall's $\tau$'s in $(-1, 1)$ via the relations in Joe (2014, Chapter 4).

Because the number of parameters is the same between the models after fixing the number of random effects, we can use the log-likelihood at the ML estimates as a rough diagnostic measure for model selection between the models. For vine copulas, Dissmann et al. (2013) found that pair-copula selection based on likelihood seems to be better than even using bivariate goodness-of-fit tests. The goodness-of-fit procedures involve a global distance measure between the model-based and empirical distribution; hence, they might not be sensitive to tail behaviors and are not diagnostic in the sense of suggesting improved parametric models in the case of small $p$-values (Joe, 2014, page 254). A larger likelihood value indicates a model that better approximates both the dependence structure of the data and the strength of dependence in the tails.

The results from fitting the trivariate 1-truncated D-vine copula mixed models with normal and beta margins are given in Table 4. The log-likelihoods showed that a trivariate 1-truncated D-vine copula mixed model with beta margins and

$$\mathrm{Cln}\{180^\circ, 270^\circ\} = \begin{cases} \text{Clayton rotated by } 180^\circ \text{ if } \tau > 0 \\ \text{Clayton rotated by } 270^\circ \text{ if } \tau < 0 \end{cases}$$

bivariate copulas provides the best fit. As a result, the estimate of overall disease prevalence is 0.588 (95% CI: 0.519-0.653), and

**TABLE 4** Maximized log-likelihoods, estimates and standard errors (SE) of the trivariate 1-truncated D-vine copula mixed models with normal and beta margins which include random effects for the prevalence $X_1$, sensitivity $X_2$, and specificity $X_4$ of the Pap test, and fixed effects for the sensitivity $\pi_3$ and specificity $\pi_5$ of the histology test (reference test).

| | BVN | | Frank | | Cln{0°,90°} | | Cln{0°,270°} | | Cln{180°,90°} | | Cln{180°,270°} | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Est. | SE | Est. | SE | Est. | SE | Est. | SE | Est. | SE |
| Normal margins | | | | | | | | | | | | |
| $\pi_1$ | 0.629 | 0.055 | 0.622 | 0.054 | 0.618 | 0.085 | 0.578 | 0.046 | 0.629 | 0.056 | 0.631 | 0.055 |
| $\pi_2$ | 0.646 | 0.047 | 0.640 | 0.046 | 0.635 | 0.050 | 0.690 | 0.046 | 0.644 | 0.048 | 0.653 | 0.054 |
| $\pi_3$ | 0.904 | 0.015 | 0.901 | 0.014 | 0.893 | 0.013 | 0.906 | 0.014 | 0.904 | 0.015 | 0.902 | 0.014 |
| $\pi_4$ | 0.836 | 0.034 | 0.846 | 0.033 | 0.856 | 0.031 | 0.825 | 0.033 | 0.835 | 0.037 | 0.843 | 0.037 |
| $\pi_5$ | 0.988 | 0.015 | 0.985 | 0.015 | 0.980 | 0.012 | 0.963 | 0.011 | 0.989 | 0.014 | 0.987 | 0.015 |
| $\sigma_1$ | 1.506 | 0.158 | 1.535 | 0.184 | 1.540 | 0.216 | 1.439 | 0.152 | 1.500 | 0.157 | 1.513 | 0.157 |
| $\sigma_2$ | 1.359 | 0.151 | 1.357 | 0.151 | 1.438 | 0.165 | 1.435 | 0.164 | 1.398 | 0.159 | 1.341 | 0.164 |
| $\sigma_4$ | 1.347 | 0.185 | 1.372 | 0.192 | 1.393 | 0.200 | 1.325 | 0.177 | 1.353 | 0.188 | 1.341 | 0.185 |
| $\tau_{12}$ | 0.073 | 0.103 | 0.059 | 0.105 | 0.000 | 0.113 | 0.000 | 0.115 | 0.105 | 0.086 | 0.099 | 0.081 |
| $\tau_{24}$ | −0.327 | 0.113 | −0.328 | 0.112 | −0.245 | 0.092 | −0.262 | 0.093 | −0.271 | 0.146 | −0.291 | 0.227 |
| $-\ell$ | 9056.26 | | 9056.27 | | 9058.19 | | 9057.92 | | 9056.96 | | 9055.35 | |
| Beta margins | | | | | | | | | | | | |
| $\pi_1$ | 0.588 | 0.034 | 0.588 | 0.035 | 0.588 | 0.034 | 0.586 | 0.036 | 0.588 | 0.034 | 0.588 | 0.034 |
| $\pi_2$ | 0.607 | 0.035 | 0.605 | 0.034 | 0.606 | 0.036 | 0.616 | 0.035 | 0.604 | 0.035 | 0.610 | 0.036 |
| $\pi_3$ | 0.901 | 0.015 | 0.898 | 0.014 | 0.900 | 0.015 | 0.898 | 0.014 | 0.903 | 0.015 | 0.901 | 0.014 |
| $\pi_4$ | 0.776 | 0.032 | 0.785 | 0.031 | 0.773 | 0.034 | 0.782 | 0.031 | 0.771 | 0.033 | 0.781 | 0.031 |
| $\pi_5$ | 0.987 | 0.017 | 0.984 | 0.016 | 0.986 | 0.016 | 0.976 | 0.012 | 0.990 | 0.015 | 0.987 | 0.016 |
| $\gamma_1$ | 0.291 | 0.034 | 0.294 | 0.034 | 0.292 | 0.034 | 0.302 | 0.033 | 0.287 | 0.034 | 0.290 | 0.033 |
| $\gamma_2$ | 0.246 | 0.036 | 0.245 | 0.035 | 0.255 | 0.036 | 0.253 | 0.037 | 0.251 | 0.035 | 0.244 | 0.036 |
| $\gamma_4$ | 0.197 | 0.039 | 0.198 | 0.040 | 0.205 | 0.042 | 0.195 | 0.039 | 0.204 | 0.041 | 0.190 | 0.037 |
| $\tau_{12}$ | 0.061 | 0.105 | 0.054 | 0.103 | 0.011 | 0.102 | 0.000 | 0.101 | 0.105 | 0.093 | 0.101 | 0.092 |
| $\tau_{24}$ | −0.312 | 0.117 | −0.309 | 0.111 | −0.277 | 0.153 | −0.275 | 0.097 | −0.288 | 0.133 | −0.293 | 0.135 |
| $-\ell$ | 9056.23 | | 9056.02 | | 9057.33 | | 9055.98 | | 9056.62 | | 9055.25 | |

The resulting model with normal margins and BVN copulas is the trivariate GLMM; $\text{Cln}\{\omega_1^\circ, \omega_2^\circ\} = \begin{cases} \text{Clayton rotated by } \omega_1^\circ & \text{if} \quad \tau > 0 \\ \text{Clayton rotated by } \omega_2^\circ & \text{if} \quad \tau < 0 \end{cases}$.

the overall sensitivity and specificity for the Pap test are estimated as 0.61 (95% CI: 0.537-0.678) and 0.781 (95% CI: 0.713-0.836), respectively. The overall sensitivity and specificity for the histology test are estimated as 0.901 (95% CI: 0.869-0.925 ) and 0.987 (95% CI: 0.869-0.999), respectively. Furthermore, the covariance between the latent sensitivity and specificity for the Pap test is estimated to be negative, as would be expected due to the trade-off between sensitivity and specificity when the cutoff value varies across studies. In contrast, in the trivariate GLMM, the estimate of overall disease prevalence is 0.629 (95% CI: 0.516-0.730), the estimated overall sensitivity and specificity for the Pap test are 0.646 (95% CI: 0.550-0.732) and 0.836 (95% CI: 0.758-0.893), and that for the histology test are 0.904 (95% CI: 0.872-0.929) and 0.988 (95% CI: 0.871-0.999), respectively. Note that the logit transformation and the delta method are used to construct the confidence intervals with input the results from Table 4.

It is revealed that a 1-truncated D-vine copula mixed model with the vector of probabilities of each combination of tests results on the original scale provides better fit than the GLMM, which models the vector of probabilities of each combination of tests results on a transformed scale. The improvement over the GLMM is small in terms of the likelihood principle, but for a sample size such as $N = 59$, $-9055.22 - (-9056.26) = 1.04$ units log-likelihood difference is sufficient.

The fact that the best-fitting bivariate copula that links the latent sensitivity and specificity of the Pap test is Clayton rotated by 270° reveals that there exists negative lower-upper tail dependence among the latent sensitivity and specificity. It is also apparent that the estimates of the meta-analytic parameters of interest from the 1-truncated D-vine copula mixed models with normal margins differentiate from the ones with beta margins. For example, the resultant meta-analytic estimate of the sensitivity of the Pap test ranges from 0.635 to 0.653 (normal margins) and from 0.604 to 0.610 (beta margins). This is consistent with the simulation results and conclusions in the upcoming section. The main parameters of interest, that is, the meta-analytic parameters of sensitivity and specificity of the Pap test, are biased when the univariate random effects are misspecified. Our general model can allow both normal and beta margins, that is, it is not restricted to normal margins as the GLMM.

## 5 SIMULATION STUDIES

An extensive simulation study was conducted to (1) examine the reliability of using the all possible random effects procedure based on information criteria to select the random effects, (2) gauge the small-sample efficiency of the proposed estimation method and investigate the misspecification of either the parametric margin or bivariate copula of the random effects distri-

bution, and (3) examine the reliability of using the likelihood to select the correct bivariate copula and margin.

In our simulations, we set the sample size, random effects, and the true univariate and dependence parameters to mimic the data on $N = 59$ studies from the meta-analysis of the Pap test that diagnoses cervical neoplasia in the preceding section. We use the following simulation process to generate data from the selected submodel C.2 in the preceding section, that is, the trivariate 1-truncated D-vine copula mixed model with random prevalence, sensitivity, and specificity of the index test and fixed sensitivity and specificity of the reference test:

(1) Simulate $(u_1, u_2, u_4)$ from a 3-variate 1-truncated D-vine distribution $C_3(\cdot; \boldsymbol{\theta})$.
(2) Convert to normal or beta realizations via $x_t = l^{-1}\left(F^{-1}(u_t; l(\pi_t), \delta_t\right), \ t = 1, 2, 4$.
(3) Simulate the study size $n$ from a shifted gamma distribution with shape of 1.2, rate of 0.01, and shift of 30 to obtain heterogeneous study sizes (Paul et al., 2010), and round off $n$ to the nearest integers.
(4) Draw $(y_{i11}, y_{i10}, y_{i01}, y_{i00})$ from $\mathcal{M}_4\left(n, p_{11}, p_{10}, p_{01}, p_{00}\right)$, where $p_{11}, p_{10}, p_{01}$ are calculated as in (8).

Findings on the reliability of the all possible random effects procedure based on information criteria are given in Web Table 1. We simulate from a trivariate 1-truncated D-vine copula model with Cln{180°,270°} copulas and normal or beta margins. The table presents the number of times each random effects model was chosen over 100 simulation runs and reveals our approach has a good probability of selecting the "true" submodel C.2. Note that the AIC tends to choose the submodel D.3 more often than BIC does because AIC is more likely to result in an overparameterized model.

Representative summaries of findings on the performance of the ML method in Section 3.3 are given in Web Tables 2 and 3 for trivariate 1-truncated D-vine copula mixed models with normal and beta margins. The true (simulated) bivariate copulas are the Cln{180°,270°}. We have estimated the 1-truncated D-vine copula mixed model with different bivariate copulas and margins. To make it easier to compare strengths of dependence among different copulas, we convert from the BVN, Frank, and (rotated) Clayton $\theta$'s to $\tau$'s via the relations in Joe (2014, Chapter 4). Web Tables 2 and 3 contain the resultant biases, root mean square errors (RMSEs), and standard deviations (SDs), along with average standard errors (ASEs) for the MLEs under different copula choices and margins. The standard errors of the MLEs are obtained via the gradients and the Hessian that were computed numerically during the maximization process.

Conclusions from the values in Web Tables 2 and 3 are the following:

(1) ML with the true 1-truncated D-vine copula mixed model is highly efficient according to the simulated biases, SDs and RMSEs.
(2) The MLEs of the random effects are not robust to margin misspecification, for example, in Web Table 2 (Web Table 3), where the true univariate margins are normal (beta), the biases for the MLEs of $\pi_4$ for the various copula mixed models with beta (normal) margins range from −0.06 (0.056) to −0.052 (0.063).
(3) The MLEs of the random effects are rather robust to bivariate copula misspecification, but their biases increase when the assumed bivariate copulas have different tail dependence behavior. For example, in Web Table 2 (Web Table 3), the biases for the MLEs of $\pi_2$ for the various copula mixed models with normal (beta) margins increase to 0.027 (0.017) when rotated Clayton copulas with opposite direction tail dependence are called.
(4) The MLEs of the variabilities of the random effects 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 Web Table 2, the bias for the MLE of $\sigma_4$ is −0.006, but it increases to 0.111 when rotated Clayton copulas with opposite direction tail dependence are called.
(5) The ML estimates of $\tau$'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, for example, in Web Table 2, the bias of $\hat{\tau}_{24}$ is −0.001 for the true 1-factor copula mixed model and 0.004 for a 1-factor copula mixed model with the true bivariate copulas but beta margins.

Finally, Web Table 4 presents the number of times each fitted model was chosen over the 10 000 simulation runs and reveals the true (simulated) model has been chosen for a considerable large number of times.

## 6 DISCUSSION

We have proposed a 1-truncated D-vine copula mixed model for meta-analysis of the accuracy of 2 diagnostic tests without a gold standard. Our model generalizes the GLMM proposed by Chu et al. (2009), which as we have shown, might lead to biased estimates of the meta-analytic parameters of interest. The improvement over the GLMM is due to the expression of the random effects distribution using a vine copula. This allows for flexible dependence modeling, which is different from assuming simple linear correlation structures and normality. The strength of multivariate meta-analysis approaches using copulas has been highlighted in the literature by Jackson and White (2018) and Jackson et al. (2020). The 1-truncated D-vine copula mixed model enables independent selection of parametric bivariate copulas and univariate margins from different parametric families. Consequently, the latent probabilities for each combination of test results can be modeled on the original proportions scale and can exhibit tail dependence.

We have developed an efficient ML estimation technique based on dependent Gauss-Legendre quadrature points, using a 1-truncated D-vine copula distribution. We utilized the concept of a truncated at level 1 D-vine copula, resulting in a significant reduction of the dependence parameters. This is highly useful for

estimating purposes, especially considering the typically small sample sizes in meta-analyses of diagnostic test accuracy studies. Chu et al. (2009) and Liu et al. (2015) estimated the GLMM using SAS PROC NLMIXED and acknowledged that optimizing the likelihood when random effects are allowed for both index and reference tests is non-trivial because it involves calculating 5-dimensional integrals numerically. Our numerical method, which is based on dependent Gauss-Legendre quadrature points that have a 1-truncated D-vine copula distribution, successively computes the 5-dimensional integrals in quintuple sums over the dependent quadrature points and weights. Nevertheless, in practice, the complexity of the models that should be considered depends on the number of studies in the meta-analysis and the degree of heterogeneity of the studies. Too many random effects with insufficient data would typically imply near non-identifiability (flat log-likelihood). We have not come across data sets with the need for more than 3 random effects, that is, submodels with fewer random effects should be considered, as done in the data analysis in Section 4.

Building on the basic model proposed in this paper, there are several extensions that can be implemented. Similarly to Wang et al. (2023), who proposed a GLMM that can incorporate study-level covariates, the 1-truncated D-vine copula mixed model can also easily incorporate study-level covariates. Furthermore, as in this article, similarly to Liu et al. (2015), we did not consider the situation where the 2 tests may be conditional dependent given the latent disease status and study-specific random effects. Further research is needed for extensions of the proposed model under such conditional dependence.

## ACKNOWLEDGMENTS

We would like to thank the co-editor, the associate editor, and one referee for comments leading to an improved presentation. The simulations presented in this paper were carried out on the High Performance Computing Cluster at the University of East Anglia.

## SUPPLEMENTARY MATERIALS

Supplementary material is available at *Biometrics* online.

Web appendices referenced in Sections 4 and 5 are available with this paper at the Biometrics website on Oxford Academic. R functions to derive estimates and simulate from the proposed models are part of the R package **CopulaREMADA** (Nikoloulopoulos, 2024a). The package is posted online with this article and also available at 10.32614/ CRAN.package.CopulaREMADA. The codes used in Section 4 are given as code examples within the package.

## FUNDING

None declared.

## CONFLICT OF INTEREST

None declared.

## DATA AVAILABILITY

The data that support the findings in this paper are available in the R package **CopulaREMADA** (Nikoloulopoulos, 2024a).

## REFERENCES

Aas, K., Czado, C., Frigessi, A. and Bakken, H. (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics & Economics*, 44, 182–198.

Barthel, N., Geerdens, C., Czado, C. and Janssen, P. (2019). Dependence modeling for recurrent event times subject to right-censoring with D-vine copulas. *Biometrics*, 75, 439–451.

Bedford, T. and Cooke, R. M. (2002). Vines—a new graphical model for dependent random variables. *Annals of Statistics*, 30, 1031–1068.

Brechmann, E. C., Czado, C. and Aas, K. (2012). Truncated regular vines in high dimensions with applications to financial data. *Canadian Journal of Statistics*, 40, 68–85.

Chu, H., Chen, S. and Louis, T. A. (2009). Random effects models in a meta-analysis of the accuracy of two diagnostic tests without a gold standard. *Journal of the American Statistical Association*, 104, 512–523.

Chu, H. and Cole, S. R. (2006). Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. *Journal of Clinical Epidemiology*, 59, 1331–1332.

Dendukuri, N., Schiller, I., Joseph, L. and Pai, M. (2012). Bayesian meta-analysis of the accuracy of a test for tuberculous pleuritis in the absence of a gold standard reference. *Biometrics*, 68, 1285–1293.

Dissmann, J., Brechmann, E., Czado, C. and Kurowicka, D. (2013). Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics & Data Analysis*, 59, 52–69.

Erhardt, T. M. and Czado, C. (2018). Standardized drought indices: a novel univariate and multivariate approach. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 67, 643–664.

Erhardt, T. M., Czado, C. and Schepsmeier, U. (2015). R-vine models for spatial time series with an application to daily mean temperature. *Biometrics*, 71, 323–332.

Fahey, M. T., Irwig, L. and Macaskill, P. (1995). Meta-analysis of pap test accuracy. *American Journal of Epidemiology*, 142, 680–689.

Gart, J. and Buck, A. (1966). Comparison of a screening test and a reference test in epidemiologic studies. II. A probabilistic model for the comparison of diagnostic tests. *American Journal of Epidemiology*, 83, 593–602.

Hoque, M. E., Acar, E. F. and Torabi, M. (2022). A time-heterogeneous D-vine copula model for unbalanced and unequally spaced longitudinal data. *Biometrics*, 79, 734–746.

Jackson, D., Riley, R. and White, I. R. (2011). Multivariate meta-analysis: potential and promise. *Statistics in Medicine*, 30, 2481–2498.

Jackson, D. and White, I. R. (2018). When should meta-analysis avoid making hidden normality assumptions? *Biometrical Journal*, 60, 1040–1058.

Jackson, D., White, I. and Riley, R. (2020). Multivariate meta-analysis. In: *Handbook of Meta-Analysis*(eds Schmid, C. H., Stijnen, T. and White, I. R.), London: Chapman & Hall.

Joe, H. (2014). *Dependence Modeling with Copulas*. London: Chapman & Hall.

Joe, H., Li, H. and Nikoloulopoulos, A. K. (2010). Tail dependence functions and vine copulas. *Journal of Multivariate Analysis*, 101, 252–270.

Kadhem, S. H. and Nikoloulopoulos, A. K. (2021). Factor copula models for mixed data. *British Journal of Mathematical and Statistical Psychology*, 74, 365–403.

Killiches, M. and Czado, C. (2018). A D-vine copula-based model for repeated measurements extending linear mixed models with homogeneous correlation structure. *Biometrics*, 74, 997–1005.

Liu, Y., Chen, Y. and Chu, H. (2015). A unification of models for meta-analysis of diagnostic accuracy studies without a gold standard. *Biometrics*, 71, 538–547.

Nikoloulopoulos, A. K. (2013). Copula-based models for multivariate discrete response data. In: *Copulae in Mathematical and Quantitative Finance*, vol. 213, (eds Durante, F., Härdle, W. and Jaworski, P.), 231–249, Berlin: Springer.

Nikoloulopoulos, A. K. (2015). 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.

Nikoloulopoulos, A. K. (2017). 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, 2270–2286.

Nikoloulopoulos, A. K. (2018a). Hybrid copula mixed models for combining case-control and cohort studies in meta-analysis of diagnostic tests. *Statistical Methods in Medical Research*, 27, 2540–2553.

Nikoloulopoulos, A. K. (2018b). On composite likelihood in bivariate meta-analysis of diagnostic test accuracy studies. *AStA Advances in Statistical Analysis*, 102, 211–227.

Nikoloulopoulos, A. K. (2019). A D-vine copula mixed model for joint meta-analysis and comparison of diagnostic tests. *Statistical Methods in Medical Research*, 28, 3286–3300.

Nikoloulopoulos, A. K. (2020a). An extended trivariate vine copula mixed model for meta-analysis of diagnostic studies in the presence of non-evaluable outcomes. *The International Journal of Biostatistics*, 16, 20190107.

Nikoloulopoulos, A. K. (2020b). A multinomial quadrivariate D-vine copula mixed model for meta-analysis of diagnostic studies in the presence of non-evaluable subjects. *Statistical Methods in Medical Research*, 29, 2988–3005.

Nikoloulopoulos, A. K. (2022). An one-factor copula mixed model for joint meta-analysis of multiple diagnostic tests. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185, 1398–1423.

Nikoloulopoulos, A. K. (2024a). *CopulaREMADA: Copula mixed models for multivariate meta-analysis of diagnostic test accuracy studies*, Vienna, Austria: R Foundation for Statistical Computing. R package version 1.7.3. 10.32614/CRAN.package.CopulaREMADA.

Nikoloulopoulos, A. K. (2024b). Joint meta-analysis of two diagnostic tests accounting for within and between studies dependence. *Statistical Methods in Medical Research*, 33, 1800–1817.

Nikoloulopoulos, A. K. and Joe, H. (2015). Factor copula models for item response data. *Psychometrika*, 80, 126–150.

Nikoloulopoulos, A. K., Joe, H. and Li, H. (2012). Vine copulas with asymmetric tail dependence and applications to financial return data. *Computational Statistics & Data Analysis*, 56, 3659–3673.

Normand, S. L. (1999). Meta-analysis: formulating, evaluating, combining, and reporting. *Statistics in Medicine*, 18, 321–59.

Paul, M., Riebler, A., Bachmann, L. M., Rue, H. and Held, L. (2010). Bayesian bivariate meta-analysis of diagnostic test studies using integrated nested laplace approximations. *Statistics in Medicine*, 29, 1325–1339.

Reitsma, J. B., Glas, A. S., Rutjes, A. W. S., Scholten, R. J. P. M., Bossuyt, P. M. and Zwinderman, A. H. (2005). Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology*, 58, 982–990.

Rutter, C. M. and Gatsonis, C. A. (2001). A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Statistics in Medicine*, 20, 2865–2884.

Sahin, O. and Czado, C. (2024). High-dimensional sparse vine copula regression with application to genomic prediction. *Biometrics*, 80, ujad042.

Wang, Z., Murray, T. A., Xiao, M., Lin, L., Alemayehu, D. and Chu, H. (2023). Bayesian hierarchical models incorporating study-level covariates for multivariate meta-analysis of diagnostic tests without a gold standard with application to COVID-19. *Statistics in Medicine*, 42, 5085–5099.



Received: July 30, 2024; Revised: January 10, 2025; Accepted: March 12, 2025