Appendix of “*Joint estimation of disease-specific sensitivities and specificities in reader-based multi-disease diagnostic studies of paired organs*”

N. Withanage^{a,b}, A. R. de Leon^{b,**}, and C. J. Rudnisky^c

^a*Department of Economics & Statistics, Sabaragamuwa University of Sri Lanka*

^b*Department of Mathematics & Statistics, University of Calgary*

^c*Department of Ophthalmology, University of Alberta*

In this appendix, we show details on the identifiability of the variance and correlation parameters of the joint model based on (23). The joint model takes the following marginal correlation matrix $\mathbf{R}_i$ with ten different marginal correlations (omitting the lower triangle for convenience):

$$\begin{matrix} & R_1P_1L & R_1P_2L & R_2P_1L & R_2P_2L & R_1P_1R & R_1P_2R & R_2P_1R & R_2P_2R \\ \begin{matrix} R_1P_1L \\ R_1P_2L \\ R_2P_1L \\ R_2P_2L \\ R_1P_1R \\ R_1P_2R \\ R_2P_1R \\ R_2P_2R \end{matrix} & \left(\begin{array}{cccccccc} 1 & r_1 & r_2 & r_3 & r_4 & r_5 & r_6 & r_7 \\ & 1 & r_3 & r_8 & r_5 & r_9 & r_7 & r_{10} \\ & & 1 & r_1 & r_6 & r_7 & r_4 & r_5 \\ & & & 1 & r_7 & r_{10} & r_5 & r_9 \\ & & & & 1 & r_1 & r_2 & r_3 \\ & & & & & 1 & r_3 & r_8 \\ & & & & & & 1 & r_1 \\ & & & & & & & 1 \end{array} \right), \end{matrix}$$

where

$$r_1 = \frac{2 + \rho^{(p)} \sqrt{\sigma_{11}^{(p)} \sigma_{22}^{(p)}}}{\sqrt{(3 + \sigma_{11}^{(p)})(3 + \sigma_{22}^{(p)})}}, \quad (1)$$

$$r_2 = \frac{1 + \rho^{(r)} + \sigma_{11}^{(p)}}{3 + \sigma_{11}^{(p)}}, \quad (2)$$

$$r_3 = \frac{1 + \rho^{(r)} + \rho^{(p)} \sqrt{\sigma_{11}^{(p)} \sigma_{22}^{(p)}}}{\sqrt{(3 + \sigma_{11}^{(p)})(3 + \sigma_{22}^{(p)})}}, \quad (3)$$

$$r_4 = \frac{1 + \rho^{(e)} + \sigma_{11}^{(p)}}{3 + \sigma_{11}^{(p)}}, \quad (4)$$

^{**}Corresponding author. Email: adeleon@ucalgary.ca

$$r_5 = \frac{1 + \rho^{(e)} + \rho^{(p)} \sqrt{\sigma_{11}^{(p)} \sigma_{22}^{(p)}}}{\sqrt{(3 + \sigma_{11}^{(p)})(3 + \sigma_{22}^{(p)})}}, \quad (5)$$

$$r_6 = \frac{\rho^{(e)} + \rho^{(r)} + \sigma_{11}^{(p)}}{3 + \sigma_{11}^{(p)}}, \quad (6)$$

$$r_7 = \frac{\rho^{(r)} + \rho^{(e)} + \rho^{(p)} \sqrt{\sigma_{11}^{(p)} \sigma_{22}^{(p)}}}{\sqrt{(3 + \sigma_{11}^{(p)})(3 + \sigma_{22}^{(p)})}}, \quad (7)$$

$$r_8 = \frac{1 + \rho^{(r)} + \sigma_{22}^{(p)}}{3 + \sigma_{22}^{(p)}}, \quad (8)$$

$$r_9 = \frac{1 + \rho^{(e)} + \sigma_{22}^{(p)}}{3 + \sigma_{22}^{(p)}}, \quad (9)$$

$$r_{10} = \frac{\rho^{(e)} + \rho^{(r)} + \sigma_{22}^{(p)}}{3 + \sigma_{22}^{(p)}}. \quad (10)$$

Note that $\mathbf{R}_i$ can be computed from the data. Following [1], the identifiability of the correlations $\rho^{(r)}$, $\rho^{(e)}$, and $\rho^{(p)}$ and the variance components $\sigma_{11}^{(p)}$ and $\sigma_{22}^{(p)}$ in the joint model is shown in the following. By (4) we have

$$\sigma_{11}^{(p)} = \frac{3r_4 - 1 - \rho^{(e)}}{1 - r_4}, \quad (11)$$

and by (2) and (6), we have

$$r_2 - r_6 = \frac{1 - \rho^{(e)}}{3 + \sigma_{11}^{(p)}}. \quad (12)$$

Now, substituting $\sigma_{11}^{(p)}$ in (11) into (12), we get

$$\rho^{(e)} = \frac{1 - r_4 - 2r_2 + 2r_6}{1 - r_4 - r_2 + r_6}. \quad (13)$$

Substituting $\rho^{(e)}$ from (13) into (11), we then get

$$\sigma_{11}^{(p)} = \frac{3r_4 + 3r_2 - 3r_6 - 2}{1 - r_4 - r_2 + r_6}. \quad (14)$$

Similarly, by (9), we have $\sigma_{22}^{(p)} = (1 + \rho^{(e)} - 3r_9)/(r_9 - 1)$, and substituting $\rho^{(e)}$ from (13), we get

$$\sigma_{22}^{(p)} = \frac{(1 - r_4)(2 - 3r_9) + 3(1 - r_9)(r_6 - r_2)}{(r_9 - 1)(1 - r_4 - r_2 + r_6)}. \quad (15)$$

Next, by (2), we have $\rho^{(r)} = (3r_2 - 1) + \sigma_{11}^{(p)}(r_2 - 1)$; from (14), we get

$$\rho^{(r)} = \frac{1 - r_2 - 2(r_4 - r_6)}{1 - r_4 - r_2 + r_6}. \quad (16)$$

Once we have $\sigma_{11}^{(p)}$ and $\sigma_{22}^{(p)}$, we can compute $\rho^{(p)}$ using (1), and hence all the variance components and correlations in the joint model are identifiable. For the regression coefficients, the intercepts are identifiable since we fixed the cutpoint to be zero; the slopes are identifiable as well since we have the true disease status.

Now, consider the following model for a single pathology:

$$Y_{ijkv}^* = \beta_{v0} + \beta_{v1}D_{ijv} + B_{i1}^{(p)} + B_{ik}^{(r)} + B_{ij}^{(e)} + \epsilon_{ijkv}^*, \quad (17)$$

with $\epsilon_{ijkv}^* \stackrel{iid}{\sim} N(0, 1)$, $B_{i1}^{(p)} \stackrel{iid}{\sim} N(0, \sigma_{11}^{(p)})$, and

$$\mathbf{B}_i^{(r)} = \begin{pmatrix} B_{i1}^{(r)} \\ B_{i2}^{(r)} \end{pmatrix} \stackrel{iid}{\sim} N_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \mathbf{\Sigma}^{(r)} = \begin{pmatrix} \sigma_{11}^{(r)} & \rho^{(r)} \sqrt{\sigma_{11}^{(r)} \sigma_{22}^{(r)}} \\ \rho^{(r)} \sqrt{\sigma_{11}^{(r)} \sigma_{22}^{(r)}} & \sigma_{22}^{(r)} \end{pmatrix} \right), \quad (18)$$

$$\mathbf{B}_i^{(e)} = \begin{pmatrix} B_{iL}^{(e)} \\ B_{iR}^{(e)} \end{pmatrix} \stackrel{iid}{\sim} N_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \mathbf{\Sigma}^{(e)} = \begin{pmatrix} 1 & \rho^{(e)} \\ \rho^{(e)} & 1 \end{pmatrix} \right). \quad (19)$$

The resulting joint model yields the following marginal correlation matrix (omitting the lower triangle for convenience):

$$\mathbf{R}_i = \begin{matrix} & R_1L & R_2L & R_1R & R_2R \\ \begin{matrix} R_1L \\ R_2L \\ R_1R \\ R_2R \end{matrix} & \begin{pmatrix} 1 & r_1 & r_2 & r_3 \\ & 1 & r_3 & r_4 \\ & & 1 & r_1 \\ & & & 1 \end{pmatrix} \end{matrix}, \quad (20)$$

where

$$\begin{aligned} r_1 &= \frac{1 + \sigma_{11}^{(p)} + \rho^{(r)} \sqrt{\sigma_{11}^{(r)} \sigma_{22}^{(r)}}}{\sqrt{(2 + \sigma_{11}^{(p)} + \sigma_{11}^{(r)})(2 + \sigma_{11}^{(p)} + \sigma_{22}^{(r)})}}, \\ r_2 &= \frac{\sigma_{11}^{(p)} + \sigma_{11}^{(r)} + \rho^{(e)}}{2 + \sigma_{11}^{(p)} + \sigma_{11}^{(r)}}, \\ r_3 &= \frac{\sigma_{11}^{(p)} + \rho^{(e)} + \rho^{(r)} \sqrt{\sigma_{11}^{(r)} \sigma_{22}^{(r)}}}{\sqrt{(2 + \sigma_{11}^{(p)} + \sigma_{11}^{(r)})(2 + \sigma_{11}^{(p)} + \sigma_{22}^{(r)})}}, \\ r_4 &= \frac{\sigma_{11}^{(p)} + \sigma_{22}^{(r)} + \rho^{(e)}}{2 + \sigma_{11}^{(p)} + \sigma_{22}^{(r)}}. \end{aligned}$$

Excluding the regression coefficients, the joint model based on (17) has five unknown parameters, namely, the correlations $\rho^{(r)}$ and $\rho^{(e)}$, and the variance components $\sigma_{11}^{(r)}$, $\sigma_{22}^{(r)}$, and $\sigma_{11}^{(p)}$. However, this model yields only four distinct marginal correlations r_1 , r_2 , r_3 , and r_4 . This implies that some parameters in the joint model are not identifiable.

Using the same argument, it also follows that

$$\mathbf{\Sigma}^{(e)} = \begin{pmatrix} \sigma_{LL}^{(e)} & \rho^{(e)} \sqrt{\sigma_{LL}^{(e)} \sigma_{RR}^{(e)}} \\ \rho^{(e)} \sqrt{\sigma_{LL}^{(e)} \sigma_{RR}^{(e)}} & \sigma_{RR}^{(e)} \end{pmatrix} \quad (21)$$

is not identifiable, as now we have more parameters than there are distinct correlations in (20).

References

- [1] J. S. Najita, Y. Li, and P. J. Catalano, *A novel application of a bivariate regression model for binary and continuous outcomes to studies of fetal toxicity*, Journal of the Royal Statistical Society: Series C 58 (2009), pp. 555–573.