U.S. FOOD & DRUG ADMINISTRATION

# REALISTIC SIMULATION OF READER STUDY DATA


XIAOCHEN ZHU[1,2] AND WEIJIE CHEN[1]

[1]DIVISION OF IMAGING, DIAGNOSTIC, AND SOFTWARE RELIABILITY, OSEL/CDRH/FDA, SILVER SPRING, MD

[2]DEPARTMENT OF STATISTICAL SCIENCE, GEORGE MASON UNIVERSITY, FAIRFAX, VA


## ABSTRACT


Multireader multicase (MRMC) studies are widely used in assessing medical imaging and computer-aided diagnosis devices to generalize the diagnostic performance of these devices to both the population of patient cases and the population of physician readers. Simulation study plays an important role in MRMC studies for validating MRMC analysis methods or sizing a pivotal study based on pilot data. A simulation model has parameters, given which data can be generated. This project concerns how to set simulation parameters such that the simulated data are realistic. We consider two types of MRMC simulation models: (1) a linear mixed-effect model to generate MRMC ROC rating data; and (2) a generalized linear mixed-effect model to generate MRMC binary data. For both models, we first derived quasi-closed-form expressions that express performance characteristics (e.g., area under the ROC curve and its U statistic variance components) as functions of simulation parameters. Based on these expressions, we then developed algorithms to convert these performance characteristic parameters to simulation model parameters. We conducted simulation studies to verify our algorithms. Using these algorithms, we analyzed a number a real-world reader study datasets to find simulation parameters that would generate data with similar performance characteristics of these real-world data.


## INTRODUCTION

**1. MRMC ROC rating data**

The linear mixed-effect model for ROC rating data [1, 2] is

$$X_{ijkt} = \overbrace{\mu_t + \tau_{it}}^{\text{fixed effects}} + \underbrace{\overbrace{R_{jt} + C_{kt} + RC_{jkt}}^{\text{not include modality}} + \overbrace{\tau R_{ijt} + \tau C_{ikt} + \tau RC_{ijkt} + E_{ijkt}}^{\text{include modality}}}_{\text{random factors}}. \quad (1)$$

- $X_{ijkt}$ is the value of the ROC for modality $i$, reader $j$, case $k$, and truth $t$
- $\{\sigma^2_{R0}, \sigma^2_{C0}, \sigma^2_{RC0}, \sigma^2_{R1}, \sigma^2_{C1}, \sigma^2_{RC1}\}$ for random terms that do not include modality.
- $\{\sigma^2_{iR0}, \sigma^2_{iC0}, \sigma^2_{iRC0}, \sigma^2_{iR1}, \sigma^2_{iC1}, \sigma^2_{iRC1}\}$ for random terms that include modality.
- $\sigma^2_{\Omega} = \sigma^2_{R0} + \sigma^2_{C0} + \sigma^2_{RC0} + \sigma^2_{R1} + \sigma^2_{C1} + \sigma^2_{RC1}$.
- $\sigma^2_{i} = \sigma^2_{iR0} + \sigma^2_{iC0} + \sigma^2_{iRC0} + \sigma^2_{iR1} + \sigma^2_{iC1} + \sigma^2_{iRC1}$, $\forall i \in \{A, B\}$.
- $V_i = var(X_{ijk'1} - X_{ijk0}) = \sigma^2_i + \sigma^2_{\Omega}$, $\forall i \in \{A, B\}$.
- $\Delta_i = \mu_1 - \mu_0 + \tau_{i1} - \tau_{i0}$, $\forall i \in \{A, B\}$.

The value of $AUC$ for modality $i$ can be expressed as

$$AUC_i = E(s(X_{ijk'1} - X_{ijk0})) = \Phi\left(\frac{\Delta_i}{\sqrt{V_i}}\right), \quad (2)$$

where $\Phi(\cdot)$ is the CDF of the Standard Normal distribution, and $s(\cdot)$ is a indicator function. The variance of $AUC$ for modality $i$ can be derived via U-statistics [3]

$$var(AUC_i) = \underline{c}_i^T \underline{M}_i, \quad Cov(AUC_A, AUC_B) = \underline{c}_{AB}^T \underline{M}_{AB}. \quad (3)$$

Therefore the variance of $(AUC_B - AUC_A)$ is given by

$$var(AUC_B - AUC_A) = var(AUC_A) + var(AUC_B) - 2 \times Cov(AUC_A, AUC_B). \quad (4)$$

**2. MRMC binary data**

The generalized linear mixed-effect model for binary data [4] is

$$X_{ijk} = \overbrace{\tau_i}^{\text{fixed effect}} + \underbrace{\overbrace{R_j + C_k + RC_{jk}}^{\text{not include modality}} + \overbrace{\tau R_{ij} + \tau C_{ik} + \tau RC_{ijk} + E_{ijk}}^{\text{include modality}}}_{\text{random factors}}, \quad (5)$$

$$s_{ijk} = \mathbf{1}_{\{X_{ijk}>0\}}. \quad (6)$$

- $s_{ijk}$ is the binary agreement determined by a comparison between reader report and reference report, for modality $i$, reader $j$, and case $k$.
- $\{\sigma^2_R, \sigma^2_C, \sigma^2_{RC}\}$ for random terms that do not include modality.
- $\{\sigma^2_{iR}, \sigma^2_{iC}, \sigma^2_{iRC}\}$ for random terms that include modality.
- $\sigma^2_{\Omega} = \sigma^2_R + \sigma^2_C + \sigma^2_{RC}$.
- $\sigma^2_i = \sigma^2_{iR} + \sigma^2_{iC} + \sigma^2_{iRC}$, $\forall i \in \{A, B\}$.
- $V_i = var(X_{ijk}) = \sigma^2_i + \sigma^2_{\Omega}$, $\forall i \in \{A, B\}$.
- $\Delta_i = \tau_i$, $\forall i \in \{A, B\}$.

Denote the agreement probability of modality $i$ as $P_i$, and denote the probability difference as $\Delta_p = P_B - P_A$. The value of the agreement probability of modality $i$ is

$$P_i = E(s_{ijk}) = E(s(X_{ijk})) = Pr(X_{ijk} \geq 0) = \Phi\left(\frac{\Delta_i}{\sqrt{V_i}}\right), \quad (7)$$

where $\Phi(\cdot)$ is the CDF of a Standard Normal distribution. The variance of agreement probability $var(P_i)$ and the variance of probability difference are

$$var(P_i) = \underline{c}_i^T \underline{M}_i, \quad Cov(P_A, P_B) = \underline{c}_{AB}^T \underline{M}_{AB}. \quad (8)$$

Therefore the variance of $(\Delta_p)$ is given by

$$var(\Delta_p) = var(P_A) + var(P_B) - 2 \times Cov(P_A, P_B). \quad (9)$$

## INVERSE PROBLEM

**From observed data to estimate simulation parameters**

Given product moments, a numerical method such as the Newton Raphson method can be used to find the variance components by solving the equations below.

$$M_{il} - \Psi\left(\frac{\Delta_i}{\sqrt{V_i}}, \frac{\Delta_i}{\sqrt{V_i}}, \frac{x}{V_i}\right) = 0, \quad M_{ABl} - \Psi\left(\frac{\Delta_A}{\sqrt{V_A}}, \frac{\Delta_B}{\sqrt{V_B}}, \frac{y}{\sqrt{V_A V_B}}\right) = 0. \quad (11)$$

**1. MRMC ROC rating data**

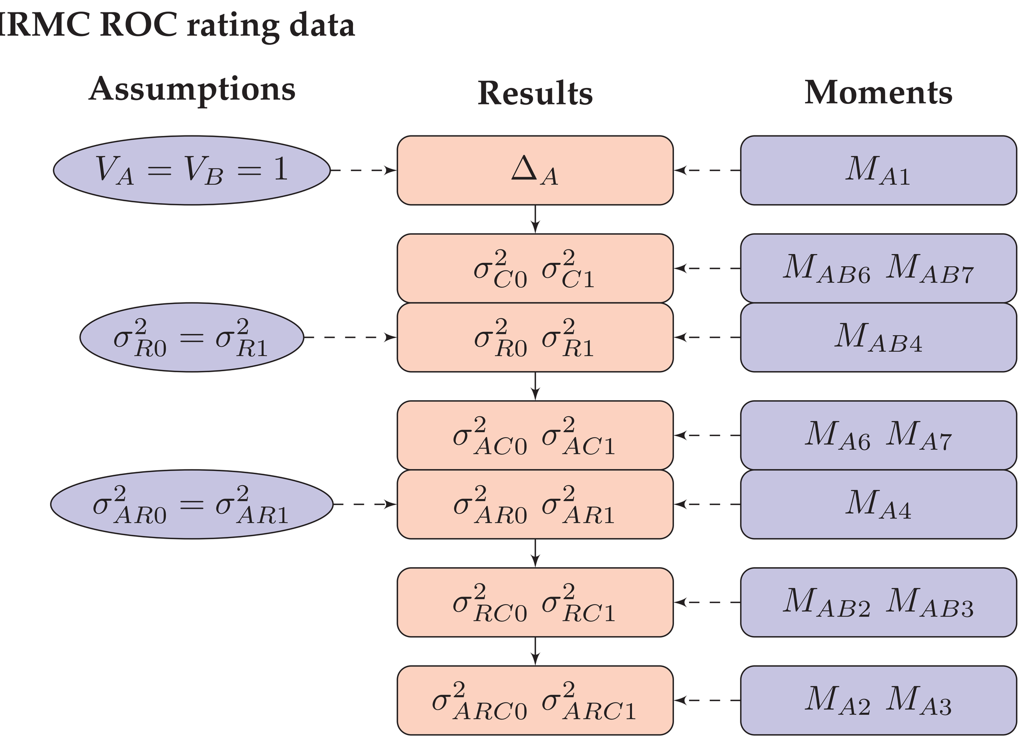


Figure 1: Flow chart of variance components estimation for MRMC ROC rating data

**2. MRMC Binary data**

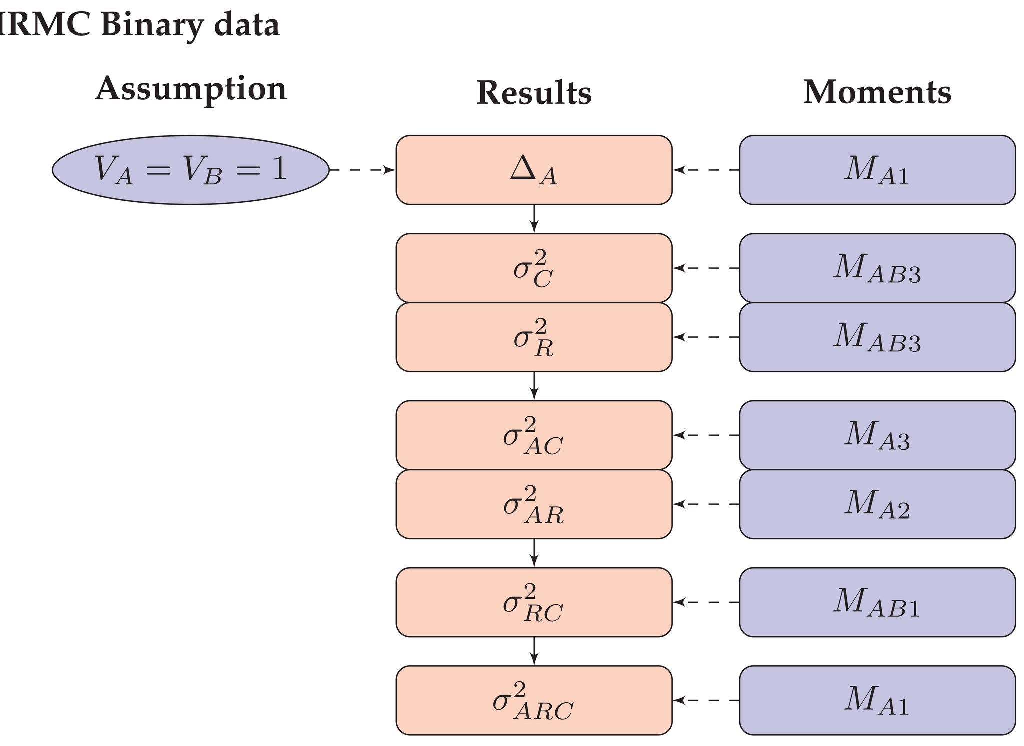


Figure 2: Flow chart of variance components estimation for MRMC binary data

## SIMULATIONS

**1. Simulation design of MRMC rating data:**

- Simulate data via (1) with:
  - $\sigma^2_{C0} = \sigma^2_{C1} = \sigma^2_{AC0} = \sigma^2_{AC1} = 0.3$,
  - $\sigma^2_{R0} = \sigma^2_{R1} = \sigma^2_{AR0} = \sigma^2_{AR1} = 0.02$,
  - $\sigma^2_{RC0} = \sigma^2_{RC1} = \sigma^2_{ARC0} = \sigma^2_{ARC1} = 0.2$.
- Size of readers and cases $(N_R, N_C)$: $(5, 1000)$ and $(10, 400)$.
- Estimate the model parameters via (11) from observed data.
- Re-simulate data via (1) with new model parameters.
- Estimate the variance of $(AUC_B - AUC_A)$ via (3) and (4) from observed data and re-simulated data.
- Iterate 10000 times for the variance components study and 1000 times for the variance of $(AUC_B - AUC_A)$ study.

**2. Simulation design of MRMC binary data:**

- Simulate data via (6) with:
  - $\sigma^2_C = \sigma^2_{AC} = 0.3$, – $\sigma^2_R = \sigma^2_{AR} = 0.02$, – $\sigma^2_{RC} = \sigma^2_{ARC} = 0.2$.
- Size of readers and cases $(N_R, N_C)$: $(5, 400)$ and $(10, 200)$.
- Estimate the model parameters via (11) from observed data.
- Re-simulate data via (6) with new model parameters.
- Estimate the variance of $(\Delta_p)$ via (8) and (9) from observed data and re-simulated data.
- Iterate 1000 times for both.

## REAL DATA STUDY

- 3 real study data sets are considered.
- Estimate the variance components from observed data via (11).
- Treat negative results as unknown values.

## FORWARD PROBLEM

**From simulation model parameters to observed performance characteristics**

Table 1: Variance components for MRMC ROC rating data

| Moments | Variance components | | | | | | |
|---|---|---|---|---|---|---|---|
| $M_{i1}$ | $\sigma^2_{i1} =$ | $\sigma^2_{iR0}$ | $+\sigma^2_{iC0}$ | $+\sigma^2_{iRC0}$ | $+\sigma^2_{iR1}$ | $+\sigma^2_{iC1}$ | $+\sigma^2_{iRC1}$ |
| $M_{i2}$ | $\sigma^2_{i2} =$ | $\sigma^2_{iR0}$ | | | $+\sigma^2_{iR1}$ | $+\sigma^2_{iC1}$ | $+\sigma^2_{iRC1}$ |
| $M_{i3}$ | $\sigma^2_{i3} =$ | $\sigma^2_{iR0}$ | $+\sigma^2_{iC0}$ | $+\sigma^2_{iRC0}$ | $+\sigma^2_{iR1}$ | | |
| $M_{i4}$ | $\sigma^2_{i4} =$ | $\sigma^2_{iR0}$ | | | $+\sigma^2_{iR1}$ | | |
| $M_{i5}$ | $\sigma^2_{i5} =$ | | $\sigma^2_{iC0}$ | | | $+\sigma^2_{iC1}$ | |
| $M_{i6}$ | $\sigma^2_{i6} =$ | | | | | $\sigma^2_{iC1}$ | |
| $M_{i7}$ | $\sigma^2_{i7} =$ | | $\sigma^2_{iC0}$ | | | | |
| $M_{i8}$ | $\sigma^2_{i8} =$ | 0 | | | | | |
| $M_{AB1}$ | $\sigma^2_{\Omega 1} =$ | $\sigma^2_{R0}$ | $+\sigma^2_{C0}$ | $+\sigma^2_{RC0}$ | $+\sigma^2_{R1}$ | $+\sigma^2_{C1}$ | $+\sigma^2_{RC1}$ |
| $M_{AB2}$ | $\sigma^2_{\Omega 2} =$ | $\sigma^2_{R0}$ | | | $+\sigma^2_{R1}$ | $+\sigma^2_{C1}$ | $+\sigma^2_{RC1}$ |
| $M_{AB3}$ | $\sigma^2_{\Omega 3} =$ | $\sigma^2_{R0}$ | $+\sigma^2_{C0}$ | $+\sigma^2_{RC0}$ | $+\sigma^2_{R1}$ | | |
| $M_{AB4}$ | $\sigma^2_{\Omega 4} =$ | $\sigma^2_{R0}$ | | | $+\sigma^2_{R1}$ | | |
| $M_{AB5}$ | $\sigma^2_{\Omega 5} =$ | | $\sigma^2_{C0}$ | | | $+\sigma^2_{C1}$ | |
| $M_{AB6}$ | $\sigma^2_{\Omega 6} =$ | | | | | $\sigma^2_{C1}$ | |
| $M_{AB7}$ | $\sigma^2_{\Omega 7} =$ | | $\sigma^2_{C0}$ | | | | |
| $M_{AB8}$ | $\sigma^2_{\Omega 8} =$ | 0 | | | | | |

Table 2: Variance components for MRMC binary data

| Moments | Variance components | | | |
|---|---|---|---|---|
| $M_{i1} = E(s_{ijk}s_{ijk})$ | $\sigma^2_{i1} =$ | $\sigma^2_{iR}$ | $+\sigma^2_{iC}$ | $+\sigma^2_{iRC}$ |
| $M_{i2} = E(s_{ijk}s_{ijk''})$ | $\sigma^2_{i2} =$ | $\sigma^2_{iR}$ | | |
| $M_{i3} = E(s_{ijk}s_{ij'k})$ | $\sigma^2_{i3} =$ | | $\sigma^2_{iC}$ | |
| $M_{i4} = E(s_{ijk}s_{ij'k'})$ | $\sigma^2_{i4} =$ | 0 | | |
| $M_{AB1} = E(s_{Ajk}s_{Bjk})$ | $\sigma^2_{\Omega 1} =$ | $\sigma^2_R$ | $+\sigma^2_C$ | $+\sigma^2_{RC}$ |
| $M_{AB2} = E(s_{Ajk}s_{Bjk'})$ | $\sigma^2_{\Omega 2} =$ | $\sigma^2_R$ | | |
| $M_{AB3} = E(s_{Ajk}s_{Bj'k})$ | $\sigma^2_{\Omega 3} =$ | | $\sigma^2_C$ | |
| $M_{AB4} = E(s_{Ajk}s_{Bj'k'})$ | $\sigma^2_{\Omega 4} =$ | 0 | | |

All product moments can be calculated using the CDF of the Standard Bivariate Normal distribution in the expression as

$$M_{il} = \Psi\left(\frac{\Delta_i}{\sqrt{V_i}}, \frac{\Delta_i}{\sqrt{V_i}}, \frac{\sigma^2_{il} + \sigma^2_{\Omega l}}{V_i}\right), \quad M_{ABl} = \Psi\left(\frac{\Delta_A}{\sqrt{V_A}}, \frac{\Delta_B}{\sqrt{V_B}}, \frac{\sigma^2_{\Omega l}}{\sqrt{V_A V_B}}\right). \quad (10)$$

where $\Psi(\cdot,\cdot,\cdot)$ is the CDF of the Standard Bivariate Normal distribution and the variance components are listed in Table 1 for MRMC ROC rating data and in Table 2 for MRMC binary data.

## RESULTS

**1. Simulation result of MRMC rating data:**

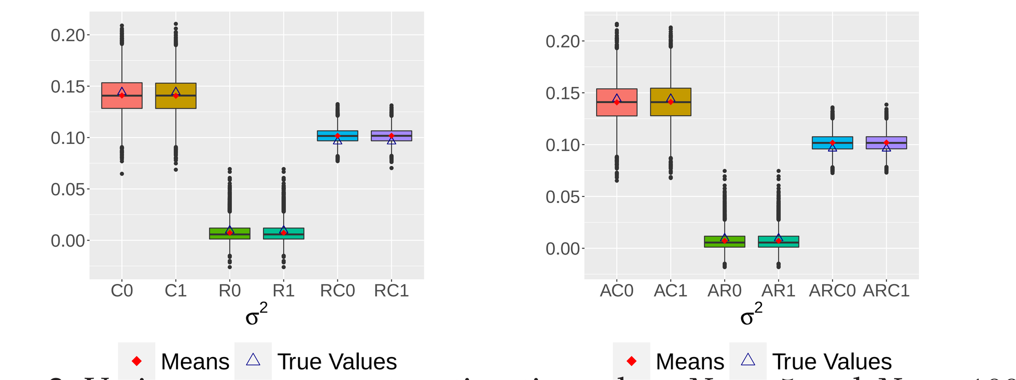


Figure 3: Variance components estimation when $N_R = 5$ and $N_C = 1000$

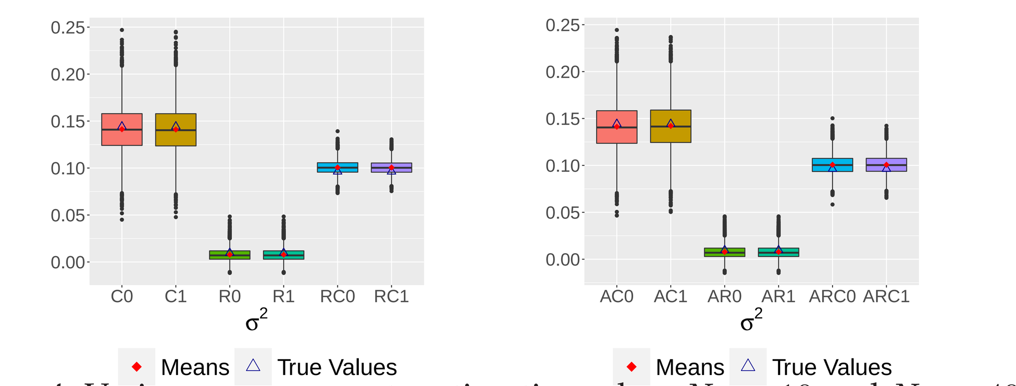


Figure 4: Variance components estimation when $N_R = 10$ and $N_C = 400$

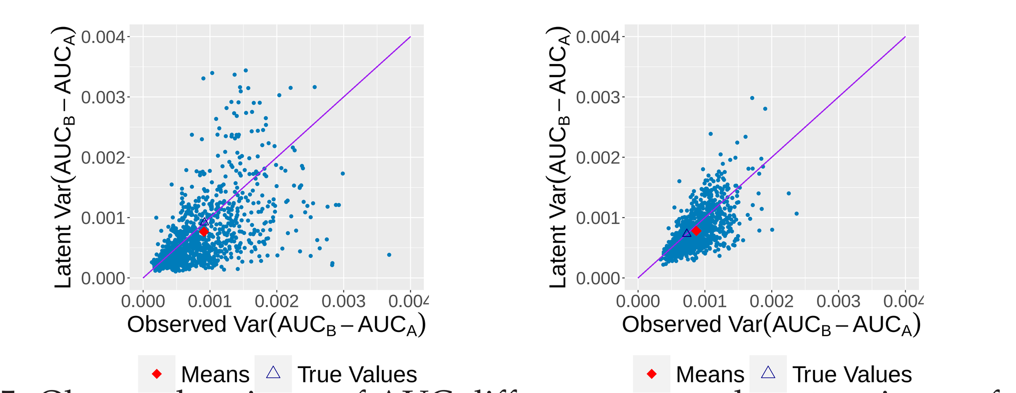


Figure 5: Observed variance of AUC difference versus latent variance of AUC difference

**2. Simulation result of MRMC binary data:**

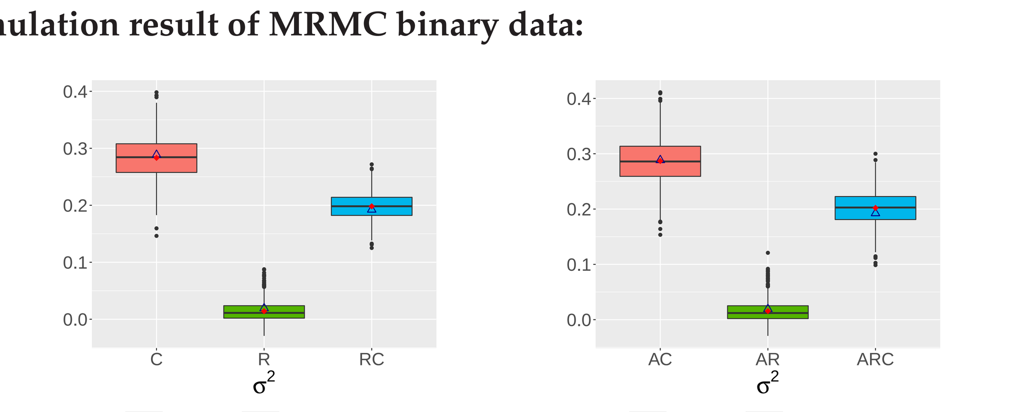


Figure 6: Variance components estimation when $N_R = 5$ and $N_C = 400$

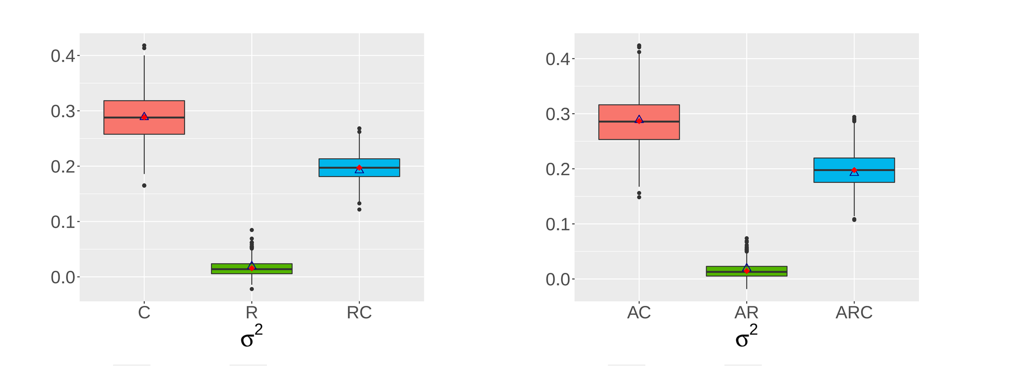


Figure 7: Variance components estimation when $N_R = 10$ and $N_C = 200$

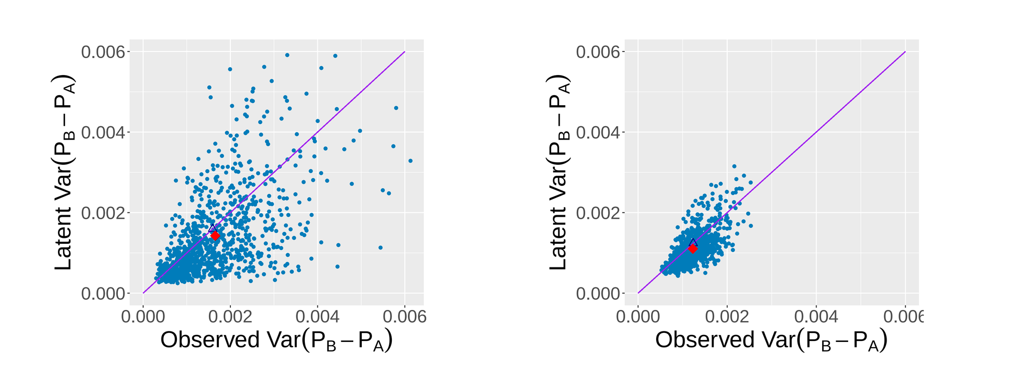


Figure 8: Observed variance of percent correct difference versus latent variance of percent correct difference

**3. Real data study:**

| $N_R$ | $N_0$ | $N_1$ | $\sigma^2_{R0}$ | $\sigma^2_{C0}$ | $\sigma^2_{RC0}$ | $\sigma^2_{R1}$ | $\sigma^2_{C1}$ | $\sigma^2_{RC1}$ |
|---|---|---|---|---|---|---|---|---|
| 10 | 32 | 35 | 0.157 | 0.344 | - | 0.157 | 0.535 | - |
| 4 | 31 | 13 | 0.337 | 0.683 | - | 0.337 | 0.856 | - |
| 14 | 260 | 48 | 0.316 | 0.656 | - | 0.316 | 0.799 | - |
| | | | $\sigma^2_{AR0}$ | $\sigma^2_{AC0}$ | $\sigma^2_{ARC0}$ | $\sigma^2_{AR1}$ | $\sigma^2_{AC1}$ | $\sigma^2_{ARC1}$ |
| 10 | 32 | 35 | - | 0.005 | 0.053 | - | 0.082 | 0.068 |
| 4 | 31 | 13 | - | - | 0.234 | - | - | 0.139 |
| 14 | 260 | 48 | - | - | 0.181 | - | - | 0.123 |
| | | | $\sigma^2_{BR0}$ | $\sigma^2_{BC0}$ | $\sigma^2_{BRC0}$ | $\sigma^2_{BR1}$ | $\sigma^2_{BC1}$ | $\sigma^2_{BRC1}$ |
| 10 | 32 | 35 | - | - | 0.121 | - | - | 0.220 |
| 4 | 31 | 13 | 0.041 | 0.068 | - | 0.041 | 0.015 | - |
| 14 | 260 | 48 | 0.042 | 0.078 | - | 0.042 | 0.071 | - |

- Dashes (-) in the table indicate that we have a negative result in the estimation of the variance component.

## CONCLUSION

Our methods are useful in estimating MRMC simulation model parameters using real-world reader study data. This would allow simulation of reader study data with performance characteristics similar to those observed in real-world experiments.

## REFERENCES


[1] Cheryl A Roe and Charles E Metz. Dorfman-berbaum-metz method for statistical analysis of multireader, multimodality receiver operating characteristic data: validation with computer simulation. *Academic radiology*, 4(4):298–303, 1997.

[2] Brandon D Gallas and Stephen L Hillis. Generalized roe and metz receiver operating characteristic model: analytic link between simulated decision scores and empirical auc variances and covariances. *Journal of Medical Imaging*, 1(3):031006, 2014.

[3] Brandon D Gallas. One-shot estimate of mrmc variance: Auc. *Academic radiology*, 13(3):353–362, 2006.

[4] Weijie Chen, Adam Wunderlich, Nicholas A Petrick, and Brandon D Gallas. Multireader multicase reader studies with binary agreement data: simulation, analysis, validation, and sizing. *Journal of Medical Imaging*, 1(3):031011, 2014.