Robust Methods for Assessment of Average and Scaled Average Bioequivalence

Robert Schall\textsuperscript{1} & Divan Aristo Burger\textsuperscript{2}

October 2019

\textsuperscript{1} Dept. of Mathematical Statistics and Actuarial Science, University of the Free State, Bloemfontein, South Africa

\textsuperscript{2} Department of Statistics, University of Pretoria, Pretoria, South Africa
Objectives

1. Propose a flexible robust methodology for bioequivalence
2. Evaluate/motivate need for robust methods in bioequivalence
3. Compare robust and conventional methods empirically:
   - Apply methodology to a reasonably large data pool of conventional bioequivalence studies
   - Apply methodology to a number of replicate design bioequivalence studies

Do the above for:
   - Conventional average bioequivalence assessment
     - Data from typical $2 \times 2$ crossover studies
   - Reference-scaled average bioequivalence assessment
     - Data from replicate design crossover studies
Objectives

1. Propose a flexible robust methodology for bioequivalence
2. Evaluate/motivate need for robust methods in bioequivalence
3. Compare robust and conventional methods empirically:
   - Apply methodology to a reasonably large data pool of conventional bioequivalence studies
   - Apply methodology to a number of replicate design bioequivalence studies

Do the above for:
- Conventional average bioequivalence assessment
  - Data from typical $2 \times 2$ crossover studies
- Reference-scaled average bioequivalence assessment
  - Data from replicate design crossover studies
Objectives

1. Propose a flexible robust methodology for bioequivalence
2. Evaluate/motivate need for robust methods in bioequivalence
3. Compare robust and conventional methods empirically:
   - Apply methodology to a reasonably large data pool of conventional bioequivalence studies
   - Apply methodology to a number of replicate design bioequivalence studies

Do the above for:

- Conventional average bioequivalence assessment
  - Data from typical $2 \times 2$ crossover studies
- Reference-scaled average bioequivalence assessment
  - Data from replicate design crossover studies
Violation of usual assumptions for “normal theory” analysis

- Outliers
- Heavy-tailed distribution (including outliers)
- Skewness of the distribution
Rationale: Approach

Replace the normal distribution by the Student $t$ distribution:
- Accommodates heavy tails/outliers
  - Small degrees of freedom
- Accommodates skewness
  - If skew Student $t$ distribution is used
  - (Only for larger datasets)

Bayesian approach:
- Noninformative priors
- Good frequentist properties
- Robust approach successfully applied even for hierarchical nonlinear models

Rationale: Approach

Replace the normal distribution by the Student $t$ distribution:
- Accommodates heavy tails/outliers
  - Small degrees of freedom
- Accommodates skewness
  - If skew Student $t$ distribution is used
  - (Only for larger datasets)

Bayesian approach:
- Noninformative priors
- Good frequentist properties
- Robust approach successfully applied even for hierarchical nonlinear models

Rationale: Approach

Replace the normal distribution by the Student $t$ distribution:

- Accommodates heavy tails/outliers
  - Small degrees of freedom
- Accommodates skewness
  - If skew Student $t$ distribution is used
  - (Only for larger datasets)

Bayesian approach:

- Noninformative priors
- Good frequentist properties
- Robust approach successfully applied even for hierarchical nonlinear models

Previous literature:

- De Souza et al. (2016):
  - Univariate and bivariate models $\Rightarrow 2 \times 2$ crossover designs
  - Extended generalized gamma distribution
  - Skew Student $t$ distribution

- Ghosh & Ntzoufras (2005):
  - Population and individual bioequivalence
  - Student $t$
Rationale: Approach

Previous literature:

- **De Souza et al. (2016):**
  - Univariate and bivariate models \( \Rightarrow 2 \times 2 \) crossover designs
  - Extended generalized gamma distribution
  - Skew Student \( t \) distribution

- **Ghosh & Ntzoufras (2005):**
  - Population and individual bioequivalence
  - Student \( t \)


Conventional model for standard $2 \times 2$ crossover study

$$y_{ij} = \mu + s_i + \zeta_h + \pi_m + \tau_j + e_{ij}$$

- $y_{ij}$ log-bioavailability for subject $i$ and formulation $j = T, R$
- $\mu$ overall mean
- $s_i$ random effects (subject)
- $\zeta_h$ (sequence), $\pi_m$ (period), $\tau_j$ (treatment): fixed effects
- $e_{ij}$ residual
- $\text{var}(s_i) = \sigma_B^2$ and $\text{var}(e_{ij}) = \sigma_W^2$
Conventional model for standard $2 \times 2$ crossover study

$$y_{ij} = \mu + s_i + \zeta_h + \pi_m + \tau_j + e_{ij}$$

- $y_{ij}$ log-bioavailability for subject $i$ and formulation $j = T, R$
- $\mu$ overall mean
- $s_i$ random effects (subject)
- $\zeta_h$ (sequence), $\pi_m$ (period), $\tau_j$ (treatment): fixed effects
- $e_{ij}$ residual
- $\text{var}(s_i) = \sigma_B^2$ and $\text{var}(e_{ij}) = \sigma_W^2$
Conventional model for standard $2 \times 2$ crossover study

$$y_{ij} = \mu + s_i + \zeta_h + \pi_m + \tau_j + e_{ij}$$

- $y_{ij}$ log-bioavailability for subject $i$ and formulation $j = T, R$
- $\mu$ overall mean
- $s_i$ random effects (subject)
- $\zeta_h$ (sequence), $\pi_m$ (period), $\tau_j$ (treatment): fixed effects
- $e_{ij}$ residual
- $\text{var}(s_i) = \sigma^2_B$ and $\text{var}(e_{ij}) = \sigma^2_W$
**Statistical decision rule**

Calculate 90% two-sided confidence interval (CI) for $\exp(\tau_T - \tau_R)$

ABE concluded if the CI falls in bioequivalence range 0.80 and 1.25
1 Subject outlier (between-subject outlier):
   - For subject $i$, both observations, $y_{iT}$ and $y_{iR}$, are extreme, in the same direction
   - Could be modeled as a mean shift in the subject effect $s_i$
   - No consequences since mean shift in $s_i$ has no effect on point or interval estimates of relative bioavailability

2 Single data point outlier (within-subject outlier):
   - For subject $i$, either $y_{iT}$ or $y_{iR}$, or both (but in opposite directions), is extreme
   - ... within-subject difference $y_{iT} - y_{iR}$ is extreme
   - Can severely affect results of the bioequivalence test

Average Bioequivalence: Types of Outlier

1. Subject outlier (between-subject outlier):
   - For subject $i$, both observations, $y_{iT}$ and $y_{iR}$, are extreme, in the same direction.
   - Could be modeled as a mean shift in the subject effect $s_i$.
   - No consequences since mean shift in $s_i$ has no effect on point or interval estimates of relative bioavailability.

2. Single data point outlier (within-subject outlier):
   - For subject $i$, either $y_{iT}$ or $y_{iR}$, or both (but in opposite directions), is extreme.
   - ... within-subject difference $y_{iT} - y_{iR}$ is extreme.
   - Can severely affect results of the bioequivalence test.

1 Subject outlier (between-subject outlier):
   - For subject $i$, both observations, $y_{iT}$ and $y_{iR}$, are extreme, in the same direction
   - Could be modeled as a mean shift in the subject effect $s_i$
   - No consequences since mean shift in $s_i$ has no effect on point or interval estimates of relative bioavailability

2 Single data point outlier (within-subject outlier):
   - For subject $i$, either $y_{iT}$ or $y_{iR}$, or both (but in opposite directions), is extreme
   - ... within-subject difference $y_{iT} - y_{iR}$ is extreme
   - Can severely affect results of the bioequivalence test

Subject outlier (between-subject outlier):
- For subject $i$, both observations, $y_{iT}$ and $y_{iR}$, are extreme, in the same direction
- Could be modeled as a mean shift in the subject effect $s_i$
- No consequences since mean shift in $s_i$ has no effect on point or interval estimates of relative bioavailability

Single data point outlier (within-subject outlier):
- For subject $i$, either $y_{iT}$ or $y_{iR}$, or both (but in opposite directions), is extreme
- ... within-subject difference $y_{iT} - y_{iR}$ is extreme
- Can severely affect results of the bioequivalence test

Robust methodology

Achieve robustness by:

1. Use standard model for $2 \times 2$ crossover study (see above)
2. But specify heavy-tailed distributions for residuals (and subject effects)
3. Here, use Student $t$ distributions with low degrees of freedom
4. (Preferable to estimate the degrees of freedom, but can also be fixed)

Fit the model using a Bayesian approach

Here: Compare Student $t$ with normal model (both Bayes)
Robust methodology

- Achieve robustness by:
  1. Use standard model for $2 \times 2$ crossover study (see above)
  2. But specify heavy-tailed distributions for residuals (and subject effects)
  3. Here, use Student $t$ distributions with low degrees of freedom
  4. (Preferable to estimate the degrees of freedom, but can also be fixed)

- Fit the model using a Bayesian approach
- Here: Compare Student $t$ with normal model (both Bayes)
Robust methodology

- Achieve robustness by:
  1. Use standard model for $2 \times 2$ crossover study (see above)
  2. But specify heavy-tailed distributions for residuals (and subject effects)
  3. Here, use Student $t$ distributions with low degrees of freedom
  4. (Preferable to estimate the degrees of freedom, but can also be fixed)

- Fit the model using a Bayesian approach

- Here: Compare Student $t$ with normal model (both Bayes)
Robust methodology

- Achieve robustness by:
  1. Use standard model for $2 \times 2$ crossover study (see above)
  2. But specify heavy-tailed distributions for residuals (and subject effects)
  3. Here, use Student $t$ distributions with low degrees of freedom
  4. (Preferable to estimate the degrees of freedom, but can also be fixed)

- Fit the model using a Bayesian approach
- **Here**: Compare Student $t$ with normal model (both Bayes)
Residuals and random effects

- **Bayes\(\mathcal{T}\)**
  - Student \( t \) distribution accommodates heavy tails/outliers
  - \( e_{ij} \sim t(0, \sigma_{W}^{2*}, v_{W}) \) where \( \sigma_{W}^{2} = \frac{v_{W}}{v_{W} - 2} \sigma_{W}^{2*} \)
  - \( s_{i} \sim t(0, \sigma_{B}^{2*}, v_{B}) \) where \( \sigma_{B}^{2} = \frac{v_{B}}{v_{B} - 2} \sigma_{B}^{2*} \)
  - \( v_{W}, v_{B} \Rightarrow \) degrees of freedom

- **Bayes\(\mathcal{N}\)**
  - Normal distribution – not robust to outliers
  - \( e_{ij} \sim \text{Normal}(0, \sigma_{W}^{2*}) \) where \( \sigma_{W}^{2} = \sigma_{W}^{2*} \)
  - \( s_{i} \sim \text{Normal}(0, \sigma_{B}^{2*}) \) where \( \sigma_{B}^{2} = \sigma_{B}^{2*} \)
Residuals and random effects

- **Bayes-$T$**
  - Student $t$ distribution accommodates heavy tails/outliers
  - ... $e_{ij} \sim t \left(0, \sigma^2_{W*}, \nu_W\right)$ where $\sigma^2_W = \frac{\nu_W}{\nu_W - 2} \sigma^2_{W*}$
  - ... $s_i \sim t \left(0, \sigma^2_{B*}, \nu_B\right)$ where $\sigma^2_B = \frac{\nu_B}{\nu_B - 2} \sigma^2_{B*}$
  - ... $\nu_W, \nu_B \Rightarrow$ degrees of freedom

- **Bayes-$N$**
  - Normal distribution – not robust to outliers
  - ... $e_{ij} \sim \text{Normal} \left(0, \sigma^2_{W*}\right)$ where $\sigma^2_W = \sigma^2_{W*}$
  - ... $s_i \sim \text{Normal} \left(0, \sigma^2_{B*}\right)$ where $\sigma^2_B = \sigma^2_{B*}$
Prior specifications

■ Vague priors

- ... $\mu, \zeta_h, \pi_m, \tau_j \sim N(0, 10000)$
- ... $\sigma_W^{-2} \sim \text{Gamma}(0.0001, 0.0001)$
- ... $\nu_W, \nu_B \sim \text{Normal}(0, 10000) \ T (2, \infty) \ (\text{half-normal})$
- ... $\sigma_B^* \sim t(0, 10000, 2) \ T (0, \infty) \ (\text{half-t})$
Implementation

- Fit model using JAGS via R package runjags
  - Student $t$: Mixture of normal & gamma distribution
    - Speeds up convergence
    - Most priors are conjugate $\Rightarrow$ fast convergence
  - ABE $\Rightarrow$ calculate:
    - Posterior estimate of $\exp(\tau_T - \tau_R)$
    - 90% highest posterior density (HPD) interval: LCL and UCL
  - Deviance information criterion (DIC): Discriminate between Bayes$_N$ & Bayes$_T$
Implementation

- Fit model using JAGS via R package runjags
- Student $t$: Mixture of normal & gamma distribution
  - Speeds up convergence
  - Most priors are conjugate $\Rightarrow$ fast convergence
- ABE $\Rightarrow$ calculate:
  - Posterior estimate of $\exp(\tau_T - \tau_R)$
  - 90% highest posterior density (HPD) interval: LCL and UCL
- Deviance information criterion (DIC): Discriminate between Bayes$_N$ & Bayes$_T$
Implementation

- Fit model using JAGS via R package runjags
- Student $t$: Mixture of normal & gamma distribution
  - Speeds up convergence
  - Most priors are conjugate $\Rightarrow$ fast convergence
- ABE $\Rightarrow$ calculate:
  - Posterior estimate of $\exp(\tau_T - \tau_R)$
  - 90% highest posterior density (HPD) interval: LCL and UCL
- Deviance information criterion (DIC): Discriminate between Bayes$_N$ & Bayes$_T$
Implementation

- Fit model using JAGS via R package runjags
- Student $t$: Mixture of normal & gamma distribution
  - Speeds up convergence
  - Most priors are conjugate $\Rightarrow$ fast convergence
- ABE $\Rightarrow$ calculate:
  - Posterior estimate of $\exp(\tau_T - \tau_R)$
  - 90% highest posterior density (HPD) interval: LCL and UCL
- Deviance information criterion (DIC): Discriminate between Bayes$_N$ & Bayes$_T$
Pool of datasets from conventional $2 \times 2$ bioequivalence studies:

- 65 datasets of T/R comparisons
- ... for both AUC and $C_{\text{max}}$

Schall (2012). The empirical coverage of confidence intervals: Point estimates and confidence intervals for confidence levels. *Biometrical J.*
Need for Robust Methodology: Empirical Study

Fit Bayes$_T$ model to each dataset in the data pool:

- Estimate degrees of freedom of Student $t$ distribution
- Compare with conventional REML & HL methods:
  - Shifts in point and interval estimates
  - Confidence interval lengths
Need for Robust Methodology: Empirical Study

Fit Bayes-$T$ model to each dataset in the data pool:

- Estimate degrees of freedom of Student $t$ distribution
- Compare with conventional REML & HL methods:
  - Shifts in point and interval estimates
  - Confidence interval lengths
Need for Robust Methodology: Degrees of Freedom

Figure: Bayes\textsubscript{T}: Estimates of Residual DF – AUC
Need for Robust Methodology: Degrees of Freedom

Figure: Bayes$_7$: Estimates of Random DF – AUC
Need for Robust Methodology: Degrees of Freedom

Figure: Bayes\textsubscript{T}: Estimates of Residual DF – $C_{\text{max}}$
Need for Robust Methodology: Degrees of Freedom

**Figure:** Bayes$_7$: Estimates of Random DF – $C_{\text{max}}$
Summary (degrees of freedom):

- 8/65 datasets for AUC and 5/65 datasets for $C_{\text{max}}$ suggest a “heavy tailed” distribution of the residuals.
- Suggests robust methodology might be needed in a small but non-negligible proportion of studies.
- (Data pool possibly biased towards “neat” datasets / “successful” studies)

- Heavy tails in subject effect distribution are rare in this data pool (homogeneous subject populations?)
- In any case irrelevant for bioequivalence assessment.
Summary (degrees of freedom):

- 8/65 datasets for AUC and 5/65 datasets for $C_{\text{max}}$ suggest a “heavy tailed” distribution of the residuals
- Suggests robust methodology might be needed in a small but non-negligible proportion of studies
- (Data pool possibly biased towards “neat” datasets / “successful” studies)

- Heavy tails in subject effect distribution are rare in this data pool (homogeneous subject populations?)
- In any case irrelevant for bioequivalence assessment
Summary (degrees of freedom):

- 8/65 datasets for AUC and 5/65 datasets for $C_{\text{max}}$ suggest a “heavy tailed” distribution of the residuals.
- Suggests robust methodology might be needed in a small but non-negligible proportion of studies.
- (Data pool possibly biased towards “neat” datasets / “successful” studies).

- Heavy tails in subject effect distribution are rare in this data pool (homogeneous subject populations?)
- In any case irrelevant for bioequivalence assessment.
Summary (degrees of freedom):

- 8/65 datasets for AUC and 5/65 datasets for $C_{\text{max}}$ suggest a “heavy tailed” distribution of the residuals
- Suggests robust methodology might be needed in a small but non-negligible proportion of studies
- (Data pool possibly biased towards “neat” datasets / “successful” studies)

- Heavy tails in subject effect distribution are rare in this data pool (homogeneous subject populations?)
- In any case irrelevant for bioequivalence assessment
Figure: Bayes$_T$ vs REML: eCDF of Shift in Estimates of ABE – AUC
Need for Robust Methodology: Shift in Point and Interval Estimates

Figure: Bayes$_T$ vs REML: eCDF of Shift in LCLs of ABE – AUC
Need for Robust Methodology: Shift in Point and Interval Estimates

**Figure:** Bayes_T vs REML: eCDF of Shift in UCLs of ABE – AUC
Figure: Bayes$_T$ vs REML: eCDF of Shift in Point Estimate of ABE – $C_{\text{max}}$
Need for Robust Methodology: Shift in Point and Interval Estimates

Figure: Bayes-$T$ vs REML: eCDF of Shift in LCL of ABE – $C_{\text{max}}$
Need for Robust Methodology: Shift in Point and Interval Estimates

Figure: Bayes$_T$ vs REML: eCDF of Shift in UCL of ABE – $C_{\text{max}}$
Summary (shift in point and interval estimates of GMR):

- In 5% of datasets the point estimate and confidence limits shift by ± 0.02 units (AUC & C_{max})

- Suggests robust methodology might be needed in a small but non-negligible proportion of studies

- (Again note possible bias towards “neat” datasets / “successful” studies)
Summary (shift in point and interval estimates of GMR):

- In 5% of datasets the point estimate and confidence limits shift by $\pm 0.02$ units ($\text{AUC \& C}_{\text{max}}$)

- Suggests robust methodology might be needed in a small but non-negligible proportion of studies

- (Again note possible bias towards “neat” datasets / “successful” studies)
Relative Confidence Interval Widths: Plot

- Relative CI widths against (estimated) residual degrees of freedom
- Relative CI widths against “need for robustness”
Relative Confidence Interval Widths: Plot

- Relative CI widths against (estimated) residual degrees of freedom
- Relative CI widths against “need for robustness”
Need for Robust Methodology: Relative Confidence Interval Widths

Figure: Bayes$_T$ vs REML: Ratio of CI Widths – AUC

[Graph showing the ratio of confidence interval widths against residual degrees of freedom.]
Need for Robust Methodology: Relative Confidence Interval Widths

Figure: Bayes$_T$ vs REML: Ratio of CI Widths – $C_{\text{max}}$
Summary (ratio of CI widths):
- Robust CIs are narrower than non-robust CIs when outliers are present (distribution is heavy-tailed)
- Robust CIs are similar to non-robust CIs when no outliers are present (distribution is not heavy-tailed)

Reason:
- Outliers/heavy tails inflate the residual variance
Summary (ratio of CI widths):

- Robust CIs are narrower than non-robust CIs when outliers are present (distribution is heavy-tailed)
- Robust CIs are similar to non-robust CIs when no outliers are present (distribution is not heavy-tailed)

Reason:

- Outliers/heavy tails inflate the residual variance
Summary (ratio of CI widths):

- Robust CIs are narrower than non-robust CIs when outliers are present (distribution is heavy-tailed)
- Robust CIs are similar to non-robust CIs when no outliers are present (distribution is not heavy-tailed)

Reason:

- Outliers/heavy tails inflate the residual variance
<table>
<thead>
<tr>
<th>Method Outcome</th>
<th>REML</th>
<th>HL</th>
<th>Bayes$_N$</th>
<th>Bayes$_T$</th>
<th>Total</th>
</tr>
</thead>
</table>
|               | Pass | Fail | Pass | Fail | Pass | Fail | Pass | Fail | n (%) |%
| REML Pass     | 117 (98.3) | 2 (1.7) | 118 (99.2) | 1 (0.8) | 118 (99.2) | 1 (0.8) | 119 (100.0) | 11 (100.0) | 119 (100.0) |
| Fail          | 4 (36.4) | 7 (63.6) | 0 (0.0) | 11 (100.0) | 1 (9.1) | 10 (90.9) | 11 (100.0) | 11 (100.0) | 11 (100.0) |
| HL Pass       | 117 (96.7) | 4 (3.3) | 116 (95.9) | 5 (4.1) | 118 (97.5) | 3 (2.5) | 121 (100.0) | 9 (100.0) | 121 (100.0) |
| Fail          | 2 (22.2) | 7 (77.8) | 2 (22.2) | 7 (77.8) | 1 (11.1) | 8 (88.9) | 1 (100.0) | 9 (100.0) | 9 (100.0) |
| Bayes$_N$ Pass| 118 (100.0) | 0 (0.0) | 116 (98.3) | 2 (1.7) | 117 (99.2) | 1 (0.8) | 118 (100.0) | 12 (100.0) | 118 (100.0) |
| Fail          | 1 (8.3) | 11 (91.7) | 5 (41.7) | 7 (58.3) | 2 (16.7) | 10 (83.3) | 1 (100.0) | 11 (100.0) | 11 (100.0) |
| Bayes$_T$ Pass| 118 (99.2) | 1 (0.8) | 118 (99.2) | 1 (0.8) | 117 (98.3) | 2 (1.7) | 119 (100.0) | 11 (100.0) | 119 (100.0) |
| Fail          | 1 (9.1) | 10 (90.9) | 3 (27.3) | 8 (72.7) | 1 (9.1) | 10 (90.9) | 1 (100.0) | 11 (100.0) | 11 (100.0) |
| Total         | 119 (91.5) | 11 (8.5) | 121 (93.1) | 9 (6.9) | 118 (90.8) | 12 (9.2) | 119 (91.5) | 11 (8.5) | 130 (100.0) |

The DIC statistic preferred Bayes$_T$ over Bayes$_N$ in 44 out of 130 cases.
Summary (method comparison):

- Bayes$_T$ method agrees well with REML analysis
- Bayes$_T$ method agrees better with REML analysis than HL method
- DIC statistic prefers Bayes$_T$ to Bayes$_N$ in about a third of cases
Summary (method comparison):

- Bayes$_T$ method agrees well with REML analysis
- Bayes$_T$ method agrees better with REML analysis than HL method

- DIC statistic prefers Bayes$_T$ to Bayes$_N$ in about a third of cases
Summary (method comparison):

- Bayes$_T$ method agrees well with REML analysis
- Bayes$_T$ method agrees better with REML analysis than HL method
- DIC statistic prefers Bayes$_T$ to Bayes$_N$ in about a third of cases
Assess REML, Bayes$_N$ & Bayes$_T$:

- Model performance
  - Bias
  - RMSE
  - Interval coverage

- Statistical power

- Parameters chosen such that ABE ratio is:
  - 1.00
  - 1.10
  - 1.20

- Data simulated for Bayes$_N$ & Bayes$_T$

- ... for Bayes$_T$: $v_W = 2.5$, $v_B = 15$ (Heavy tailed)
Assess REML, Bayes $N$ & Bayes $T$:

- Model performance
  - Bias
  - RMSE
  - Interval coverage

- Statistical power
  - Parameters chosen such that ABE ratio is:
    - 1.00
    - 1.10
    - 1.20

- Data simulated for Bayes $N$ & Bayes $T$
- ... for Bayes $T$: $\nu_W = 2.5, \nu_B = 15$ (Heavy tailed)
Assess REML, Bayes_N & Bayes_T:
- Model performance
  - Bias
  - RMSE
  - Interval coverage
- Statistical power
- Parameters chosen such that ABE ratio is:
  - 1.00
  - 1.10
  - 1.20
- Data simulated for Bayes_N & Bayes_T
- ... for Bayes_T: v_W = 2.5, v_B = 15 (Heavy tailed)
Assess REML, Bayes$_N$ & Bayes$_T$:

- Model performance
  - Bias
  - RMSE
  - Interval coverage

- Statistical power

- Parameters chosen such that ABE ratio is:
  - 1.00
  - 1.10
  - 1.20

- Data simulated for Bayes$_N$ & Bayes$_T$

- ... for Bayes$_T$: $\nu_W = 2.5$, $\nu_B = 15$ (Heavy tailed)
### Table: Simulation Study: Data from Normal Distribution

<table>
<thead>
<tr>
<th>ABE</th>
<th>Method</th>
<th>Bias</th>
<th>RMSE</th>
<th>Coverage</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>REML</td>
<td>0.0001</td>
<td>0.0199</td>
<td>89.9</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td>Bayes$_N$</td>
<td>0.0001</td>
<td>0.0199</td>
<td>89.5</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td>Bayes$_T$</td>
<td>0.0001</td>
<td>0.0199</td>
<td>89.6</td>
<td>100.0</td>
</tr>
<tr>
<td>1.10</td>
<td>REML</td>
<td>−0.0003</td>
<td>0.0221</td>
<td>89.4</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td>Bayes$_N$</td>
<td>−0.0003</td>
<td>0.0221</td>
<td>89.4</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td>Bayes$_T$</td>
<td>−0.0003</td>
<td>0.0221</td>
<td>89.3</td>
<td>100.0</td>
</tr>
<tr>
<td>1.20</td>
<td>REML</td>
<td>0.0002</td>
<td>0.0242</td>
<td>89.9</td>
<td>63.5</td>
</tr>
<tr>
<td></td>
<td>Bayes$_N$</td>
<td>0.0002</td>
<td>0.0242</td>
<td>90.0</td>
<td>64.1</td>
</tr>
<tr>
<td></td>
<td>Bayes$_T$</td>
<td>0.0001</td>
<td>0.0243</td>
<td>89.6</td>
<td>64.4</td>
</tr>
</tbody>
</table>
**Table:** Simulation Study: Data from Student $t$ Distribution

<table>
<thead>
<tr>
<th>ABE</th>
<th>Method</th>
<th>Bias</th>
<th>RMSE</th>
<th>Coverage</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>REML</td>
<td>−0.0007</td>
<td>0.0404</td>
<td>91.0</td>
<td>98.6</td>
</tr>
<tr>
<td></td>
<td>Bayes$_N$</td>
<td>−0.0007</td>
<td>0.0404</td>
<td>91.3</td>
<td>98.7</td>
</tr>
<tr>
<td></td>
<td>Bayes$_T$</td>
<td>−0.0001</td>
<td>0.0290</td>
<td>92.3</td>
<td>100.0</td>
</tr>
<tr>
<td>1.10</td>
<td>REML</td>
<td>−0.0002</td>
<td>0.0524</td>
<td>89.5</td>
<td>93.1</td>
</tr>
<tr>
<td></td>
<td>Bayes$_N$</td>
<td>−0.0002</td>
<td>0.0524</td>
<td>89.8</td>
<td>93.4</td>
</tr>
<tr>
<td></td>
<td>Bayes$_T$</td>
<td>−0.0017</td>
<td>0.0339</td>
<td>90.0</td>
<td>99.4</td>
</tr>
<tr>
<td>1.20</td>
<td>REML</td>
<td>0.0016</td>
<td>0.0515</td>
<td>90.8</td>
<td>29.6</td>
</tr>
<tr>
<td></td>
<td>Bayes$_N$</td>
<td>0.0016</td>
<td>0.0515</td>
<td>91.3</td>
<td>31.2</td>
</tr>
<tr>
<td></td>
<td>Bayes$_T$</td>
<td>0.0006</td>
<td>0.0358</td>
<td>91.7</td>
<td>36.3</td>
</tr>
</tbody>
</table>
Simulation Study

Summary:

- Data simulated from Bayes_N
  - All models have good accuracy, precision & coverage
  - All models yield similar statistical power

- Data simulated from Bayes_T
  - All models have good accuracy & coverage
  - Bayes_T has better precision
  - Bayes_T yields higher statistical power
Summary:

- Data simulated from Bayes$_N$
- ... All models have good accuracy, precision & coverage
- ... All models yield similar statistical power

- Data simulated from Bayes$_T$
- ... All models have good accuracy & coverage
- ... Bayes$_T$ has better precision
- ... Bayes$_T$ yields higher statistical power
Conventional model for 2-treatment, 2-sequence, 4-period replicate crossover study

\[ y_{ijk} = \mu + \zeta_h + s_{ij} + \pi_m + \tau_j + e_{ijk} \]

- \( y_{ijk} \) log-bioavailability for the \( k^{th} \) replicate of subject \( i \) and formulation \( j = T, R; k = 1, 2 \)
- \( \mu \) overall mean
- \( s_{ij} \) random effect of subject \( i \) and formulation \( j \)
- \( \zeta_h \) (sequence), \( \pi_m \) (period), \( \tau_j \) (treatment): fixed effects
- \( e_{ijk} \) residual
Scaled Average Bioequivalence: Statistical Model

Conventional model for 2-treatment, 2-sequence, 4-period replicate crossover study

\[ y_{ijk} = \mu + \zeta_h + s_{ij} + \pi_m + \tau_j + e_{ijk} \]

- \( y_{ijk} \): log-bioavailability for the \( k^{th} \) replicate of subject \( i \) and formulation \( j = T, R; \ k = 1, 2 \)
- \( \mu \): overall mean
- \( s_{ij} \): random effect of subject \( i \) and formulation \( j \)
- \( \zeta_h \): (sequence), \( \pi_m \): (period), \( \tau_j \): (treatment): fixed effects
- \( e_{ijk} \): residual
Variance-covariance matrix of the $s_i = (s_{iR}, s_{iT})'$:

$$\text{cov}(s_i) = \Sigma_B = \begin{pmatrix} \sigma^2_{BR} & \rho \sigma_{BR} \sigma_{BT} \\ \rho \sigma_{BR} \sigma_{BT} & \sigma^2_{BT} \end{pmatrix}$$

- **Between-subject variances:**
  - $\text{var}(s_{iR}) = \sigma^2_{BR}$
  - $\text{var}(s_{iT}) = \sigma^2_{BT}$
  - $\sigma^2_D = \text{var}(s_{iR} - s_{iT}) = (\sigma^2_{BR} + \sigma^2_{BT} - 2\rho \sigma_{BR} \sigma_{BT})$: subject-by-formulation interaction variance

- **Within-subject variances:**
  - $\text{var}(e_{iRk}) = \sigma^2_{WR}$
  - $\text{var}(e_{iTk}) = \sigma^2_{WT}$
Scaled Average Bioequivalence: Statistical Model

- Variance-covariance matrix of the $s_i = (s_{iR}, s_{iT})'$:

$$\text{cov}(s_i) = \Sigma_B = \begin{pmatrix} 
\sigma_{BR}^2 & \rho \sigma_{BR} \sigma_{BT} \\
\rho \sigma_{BR} \sigma_{BT} & \sigma_{BT}^2 
\end{pmatrix}$$

- Between-subject variances:
  - $\text{var}(s_{iR}) = \sigma_{BR}^2$
  - $\text{var}(s_{iT}) = \sigma_{BT}^2$
  - $\sigma_D^2 = \text{var}(s_{iR} - s_{iT}) = (\sigma_{BR}^2 + \sigma_{BT}^2 - 2\rho \sigma_{BR} \sigma_{BT})$: subject-by-formulation interaction variance

- Within-subject variances:
  - $\text{var}(e_{iRk}) = \sigma_{WR}^2$
  - $\text{var}(e_{iT_k}) = \sigma_{WT}^2$
Scaled Average Bioequivalence: Statistical Model

- Variance-covariance matrix of the \( s_i = (s_{iR}, s_{iT})' \):

\[
\text{cov}(s_i) = \Sigma_B = \begin{pmatrix}
\sigma_{BR}^2 & \rho \sigma_{BR} \sigma_{BT} \\
\rho \sigma_{BR} \sigma_{BT} & \sigma_{BT}^2
\end{pmatrix}
\]

- Between-subject variances:
  - \( \text{var}(s_{iR}) = \sigma_{BR}^2 \)
  - \( \text{var}(s_{iT}) = \sigma_{BT}^2 \)
  - \( \sigma_D^2 = \text{var}(s_{iR} - s_{iT}) = (\sigma_{BR}^2 + \sigma_{BT}^2 - 2\rho \sigma_{BR} \sigma_{BT}) \): subject-by-formulation interaction variance

- Within-subject variances:
  - \( \text{var}(e_{iRk}) = \sigma_{WR}^2 \)
  - \( \text{var}(e_{iTk}) = \sigma_{WT}^2 \)
Linearized reference-scaled average bioequivalence (RSABE) criterion:

\[ \theta = \tau^2 - k^2 \sigma_{WR}^2 \]

where

- \( k = \frac{\log(1.25)}{\sigma_0} \)
- Choose: \( \sigma_0 = 0.25 \)
Statistical decision rule

Two formulations are bioequivalent if the one-sided 95% upper confidence limit for $\theta$ is below zero.
1. **Subject outlier (between-subject outlier):**
   - For subject \( i \), all 4 observations \( y_{iT1}, y_{iT2}, y_{iR1}, y_{iR2} \) are extreme, in the same direction.
   - Could be modeled as a mean shift in \( s_i \).
   - Again: has no consequences since a mean shift in \( s_i \) does not affect point or interval estimate of relative bioavailability.

2. **Subject-by-formulation outlier (within-subject but between-replicate outlier):**
   - For subject \( i \), the replicates \( y_{iT1} \) and \( y_{iT2} \), jointly, are different from the replicates \( y_{iR1} \) and \( y_{iR2} \).
   - For subject \( i \) the within-subject, between replicate difference \((y_{iT1} + y_{iT2}) - (y_{iR1} + y_{iR2})\) is extreme.
   - Could be modeled as a mean shift in either \( s_{iT} \) or \( s_{iR} \).
   - Can severely affect results of the bioequivalence test.
Scaled Average Bioequivalence: Types of Outlier

1. Subject outlier (between-subject outlier):
   - For subject $i$, all 4 observations $y_{iT1}, y_{iT2}, y_{iR1}, y_{iR2}$ are extreme, in the same direction
   - Could be modeled as a mean shift in $s_i$
   - Again: has no consequences since a mean shift in $s_i$ does not affect point or interval estimate of relative bioavailability

2. Subject-by-formulation outlier (within-subject but between-replicate outlier):
   - For subject $i$, the replicates $y_{iT1}$ and $y_{iT2}$, jointly, are different from the replicates $y_{iR1}$ and $y_{iR2}$
   - For subject $i$ the within-subject, between replicate difference $(y_{iT1} + y_{iT2}) - (y_{iR1} + y_{iR2})$ is extreme
   - Could be modeled as a mean shift in either $s_{iT}$ or $s_{iR}$
   - Can severely affect results of the bioequivalence test
Subject outlier (between-subject outlier):
- For subject $i$, all 4 observations $y_{iT1}$, $y_{iT2}$, $y_{iR1}$, $y_{iR2}$ are extreme, in the same direction
- Could be modeled as a mean shift in $s_i$
- Again: has no consequences since a mean shift in $s_i$ does not affect point or interval estimate of relative bioavailability

Subject-by-formulation outlier (within-subject but between-replicate outlier):
- For subject $i$, the replicates $y_{iT1}$ and $y_{iT2}$, jointly, are different from the replicates $y_{iR1}$ and $y_{iR2}$
- For subject $i$ the within-subject, between replicate difference $(y_{iT1} + y_{iT2}) - (y_{iR1} + y_{iR2})$ is extreme
- Could be modeled as a mean shift in either $s_{iT}$ or $s_{iR}$
- Can severely affect results of the bioequivalence test
1  Subject outlier (between-subject outlier):
   - For subject $i$, all 4 observations $y_{iT1}$, $y_{iT2}$, $y_{iR1}$, $y_{iR2}$ are extreme, in the same direction
   - Could be modeled as a mean shift in $s_i$
   - Again: has no consequences since a mean shift in $s_i$ does not affect point or interval estimate of relative bioavailability

2  Subject-by-formulation outlier (within-subject but between-replicate outlier):
   - For subject $i$, the replicates $y_{iT1}$ and $y_{iT2}$, jointly, are different from the replicates $y_{iR1}$ and $y_{iR2}$
   - For subject $i$ the within-subject, between replicate difference $(y_{iT1} + y_{iT2}) - (y_{iR1} + y_{iR2})$ is extreme
   - Could be modeled as a mean shift in either $s_{iT}$ or $s_{iR}$
   - Can severely affect results of the bioequivalence test
1. **Subject outlier (between-subject outlier):**
   - For subject \( i \), all 4 observations \( y_{iT1}, y_{iT2}, y_{iR1}, y_{iR2} \) are extreme, in the same direction.
   - Could be modeled as a mean shift in \( s_i \).
   - Again: has no consequences since a mean shift in \( s_i \) does not affect point or interval estimate of relative bioavailability.

2. **Subject-by-formulation outlier (within-subject but between-replicate outlier):**
   - For subject \( i \), the replicates \( y_{iT1} \) and \( y_{iT2} \), jointly, are different from the replicates \( y_{iR1} \) and \( y_{iR2} \).
   - For subject \( i \) the within-subject, between replicate difference \( (y_{iT1} + y_{iT2}) - (y_{iR1} + y_{iR2}) \) is extreme.
   - Could be modeled as a mean shift in either \( s_{iT} \) or \( s_{iR} \).
   - Can severely affect results of the bioequivalence test.
3 Single data point outlier (within-subject, within-replicate outlier):

- For subject $i$, one of the four observations $y_{iT1}$, $y_{iT2}$, $y_{iR1}$, or $y_{iR2}$ is extreme
- At least one of the within-subject, within-replicate differences $(y_{iT1} - y_{iT2})$ or $(y_{iR1} - y_{iR2})$ is extreme
- Can severely affect the results of the bioequivalence test

3 Single data point outlier (within-subject, within-replicate outlier):

- For subject $i$, one of the four observations $y_{iT1}$, $y_{iT2}$, $y_{iR1}$, or $y_{iR2}$ is extreme
- At least one of the within-subject, within-replicate differences $(y_{iT1} - y_{iT2})$ or $(y_{iR1} - y_{iR2})$ is extreme
- Can severely affect the results of the bioequivalence test

Figure: Study 1a – Conditional Residuals: AUC (Slide 58)
Figure: Study 1a: $s_{DT}$ – Residuals: AUC
Figure: Study 1a: $s_{DR} -$ Residuals: AUC
Figure: Study 1a: $s_{SF}$ – Residuals: AUC
Figure: Study 1a: $s_S$ – Residuals: AUC
Scaled Average Bioequivalence: QQ and Residual Plots

**Figure:** Study 7 – Conditional Residuals: $C_{\text{max}}$ (Slide 58)
Figure: Study 7: $s_{DT}$ – Residuals: $C_{\text{max}}$
Figure: Study 7: $s_{DR}$ – Residuals: $C_{max}$
Figure: Study 7: $s_{SF}$ – Residuals: $C_{max}$
Figure: Study 7: \( s_s \) – Residuals: \( C_{\text{max}} \)
Robust methodology

- Achieve robustness by:
  1. Use standard model for replicate-design crossover study (see above)
  2. But specify heavy-tailed distributions for residuals (and subject effects)
  3. Here, use Student $t$ distributions with low degrees of freedom
  4. (Again: Preferable to estimate the degrees of freedom)

- Fit the model using a Bayesian approach
- Compare Student $t$ with normal model (both Bayes)
Robust methodology

- Achieve robustness by:
  1. Use standard model for replicate-design crossover study (see above)
  2. But specify heavy-tailed distributions for residuals (and subject effects)
  3. Here, use Student $t$ distributions with low degrees of freedom
  4. (Again: Preferable to estimate the degrees of freedom)

- Fit the model using a Bayesian approach
- Compare Student $t$ with normal model (both Bayes)
Robust methodology

- Achieve robustness by:
  1. Use standard model for replicate-design crossover study (see above)
  2. But specify heavy-tailed distributions for residuals (and subject effects)
  3. Here, use Student $t$ distributions with low degrees of freedom
  4. (Again: Preferable to estimate the degrees of freedom)

- Fit the model using a Bayesian approach
- Compare Student $t$ with normal model (both Bayes)
Robust methodology

- Achieve robustness by:
  1. Use standard model for replicate-design crossover study (see above)
  2. But specify heavy-tailed distributions for residuals (and subject effects)
  3. Here, use Student $t$ distributions with low degrees of freedom
  4. (Again: Preferable to estimate the degrees of freedom)

- Fit the model using a Bayesian approach
- Compare Student $t$ with normal model (both Bayes)
Residuals and random effects

- **Bayes\(_T\)**
  - Student \(t\) distribution accommodates heavy tails/outliers
  - \( e_{ijk} \sim t\left(0, \sigma_{Wj}^2, v_{Wj}\right) \) where \( \sigma_{Wj}^2 = \frac{v_{Wj}}{v_{Wj} - 2} \sigma_{Wj}^2 \)
  - \( s_i \sim t\left(0, \Sigma_B^*, v_B\right) \) where \( \Sigma_B = \frac{v_B}{v_B - 2} \Sigma_B^* \)
  - \( v_{Wj}, v_B \Rightarrow \) degrees of freedom

- **Bayes\(_N\)**
  - Normal distribution – not robust to outliers
  - \( e_{ijk} \sim \text{Normal}\left(0, \sigma_{Wj}^2\right) \) where \( \sigma_{Wj}^2 = \sigma_{Wj}^2 \)
  - \( s_i \sim \text{Normal}\left(0, \Sigma_B^*, v_B\right) \) where \( \Sigma_B = \Sigma_B^* \)
Residuals and random effects

Bayes$_T$

- Student $t$ distribution accommodates heavy tails/outliers
  
  \[
  e_{ijk} \sim t \left(0, \sigma_{Wj}^2, \nu_{Wj}\right) \quad \text{where} \quad \sigma_{Wj}^2 = \frac{\nu_{Wj}}{\nu_{Wj} - 2} \sigma_{Wj}^{2*}
  \]
  
  \[
  s_i \sim t \left(0, \Sigma_B^*, \nu_B\right) \quad \text{where} \quad \Sigma_B = \frac{\nu_B}{\nu_B - 2} \Sigma_B^*
  \]
  
  \[
  \nu_{Wj}, \nu_B \Rightarrow \text{degrees of freedom}
  \]

Bayes$_N$

- Normal distribution – not robust to outliers
  
  \[
  e_{ijk} \sim \text{Normal} \left(0, \sigma_{Wj}^{2*}\right) \quad \text{where} \quad \sigma_{Wj}^2 = \sigma_{Wj}^{2*}
  \]
  
  \[
  s_i \sim \text{Normal} \left(0, \Sigma_B^*, \nu_B\right) \quad \text{where} \quad \Sigma_B = \Sigma_B^*
  \]
Prior specification

- Vague priors
- ... $\mu, \zeta_h, \pi_m, \tau_j \sim N(0, 10000)$
- ... $\sigma_{Wj}^{-2*} \sim \text{Gamma}(0.0001, 0.0001)$
- ... $\nu_{Wj}, \nu_B \sim \text{Normal}(0, 10000) \ T(2, \infty)$ (half-normal)
- ... $\Sigma^* \sim \text{MGH-}t(10000, 2)$ (matrix generalized half-$t$)
Implementation

- Fit model using JAGS via R package runjags
- Student \( t \): Mixture of normal & gamma distribution
  - Speeds up convergence
  - Matrix generalized half-\( t \) for covariance matrix: Huang & Wand (2013)
  - Most priors are conjugate \( \Rightarrow \) fast convergence
- RSABE \( \Rightarrow \) calculate:
  - Posterior estimate of \( \theta \)
  - Upper limit of one-sided 95% Bayesian credibility (BCI) interval for \( \theta \)
- DIC statistic: Discriminate between Bayes\( _{N} \) & Bayes\( _{T} \)

Implementation

- Fit model using JAGS via R package runjags
- Student $t$: Mixture of normal & gamma distribution
  - Speeds up convergence
  - Matrix generalized half-$t$ for covariance matrix: Huang & Wand (2013)
  - Most priors are conjugate $\Rightarrow$ fast convergence
- RSABE $\Rightarrow$ calculate:
  - Posterior estimate of $\theta$
  - Upper limit of one-sided 95% Bayesian credibility (BCI) interval for $\theta$
- DIC statistic: Discriminate between Bayes$_N$ & Bayes$_T$

Implementation

- Fit model using JAGS via R package runjags
- Student $t$: Mixture of normal & gamma distribution
  - Speeds up convergence
  - Matrix generalized half-$t$ for covariance matrix: Huang & Wand (2013)
  - Most priors are conjugate $\Rightarrow$ fast convergence
- RSABE $\Rightarrow$ calculate:
  - Posterior estimate of $\theta$
  - Upper limit of one-sided 95% Bayesian credibility (BCI) interval for $\theta$
- DIC statistic: Discriminate between Bayes$_N$ & Bayes$_T$

Implementation

- Fit model using JAGS via R package runjags
- Student $t$: Mixture of normal & gamma distribution
  - Speeds up convergence
  - Matrix generalized half-$t$ for covariance matrix: Huang & Wand (2013)
  - Most priors are conjugate $\Rightarrow$ fast convergence
- RSABE $\Rightarrow$ calculate:
  - Posterior estimate of $\theta$
  - Upper limit of one-sided 95% Bayesian credibility (BCI) interval for $\theta$
- DIC statistic: Discriminate between Bayes$_N$ & Bayes$_T$

Pool of datasets from 2-sequence, 2-treatment, 4-period replicate designs:

- 7 studies
- ... for both AUC and $C_{\text{max}}$
- ... Study 1a: Schall, Ring, Endrenyi (2010)
- ... Study 1a: R contains outliers, whereas T does not
- ... Study 1b: Study 1a’s T and R labels reversed
### Table: AUC and $C_{\text{max}}$: UCLs of RSABE and Estimates of DF

<table>
<thead>
<tr>
<th>Study</th>
<th>Parameter</th>
<th>UCL of RSABE</th>
<th>Estimates of DF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>REML Bayes$_N$ Bayes$_T$</td>
<td>Residual Test</td>
</tr>
<tr>
<td>1a</td>
<td>AUC†</td>
<td>-0.0294 -0.0270 -0.0126</td>
<td>71.6</td>
</tr>
<tr>
<td></td>
<td>$C_{\text{max}}$†</td>
<td>-0.0262 -0.0155 -0.0108</td>
<td>68.3</td>
</tr>
<tr>
<td>1b</td>
<td>AUC†</td>
<td>-0.0091 -0.0065 -0.0107</td>
<td>3.2</td>
</tr>
<tr>
<td></td>
<td>$C_{\text{max}}$†</td>
<td>-0.0190 -0.0087 -0.0201</td>
<td>4.0</td>
</tr>
<tr>
<td>2</td>
<td>AUC†</td>
<td>-0.1388 -0.1283 -0.1245</td>
<td>39.0</td>
</tr>
<tr>
<td></td>
<td>$C_{\text{max}}$†</td>
<td>-0.1328 -0.1249 -0.1191</td>
<td>18.5</td>
</tr>
<tr>
<td>3</td>
<td>AUC†</td>
<td>-0.1382 -0.1295 -0.1228</td>
<td>17.8</td>
</tr>
<tr>
<td></td>
<td>$C_{\text{max}}$†</td>
<td>-0.1321 -0.1251 -0.1204</td>
<td>31.8</td>
</tr>
<tr>
<td>4</td>
<td>AUC†</td>
<td>-0.1757 -0.1572 -0.1510</td>
<td>38.3</td>
</tr>
<tr>
<td></td>
<td>$C_{\text{max}}$†</td>
<td>-0.2234 -0.2003 -0.1926</td>
<td>31.2</td>
</tr>
<tr>
<td>5</td>
<td>AUC†</td>
<td>-0.0085 -0.0066 -0.0065</td>
<td>13.1</td>
</tr>
<tr>
<td></td>
<td>$C_{\text{max}}$†</td>
<td>-0.0070 -0.0055 -0.0059</td>
<td>7.9</td>
</tr>
<tr>
<td>6</td>
<td>AUC†</td>
<td>-0.0013 0.0009 -0.0007</td>
<td>29.8</td>
</tr>
<tr>
<td></td>
<td>$C_{\text{max}}$†</td>
<td>-0.0009 0.0011 0.0009</td>
<td>28.5</td>
</tr>
<tr>
<td>7</td>
<td>AUC</td>
<td>-0.0517 -0.0471 -0.0474</td>
<td>62.6</td>
</tr>
<tr>
<td></td>
<td>$C_{\text{max}}$</td>
<td>0.0821 0.1055 0.1060</td>
<td>79.1</td>
</tr>
</tbody>
</table>

†DIC statistic prefers Bayes$_T$ over Bayes$_N$. 

104 / 111
Summary (need for robust methodology):

- Degrees of freedom estimates for 2/8 studies (R treatment!) are small for both AUC and $C_{\text{max}}$
- DIC prefers Bayes$_T$ to Bayes$_N$ in most cases
- Bayesian CIs are generally somewhat wider than REML CIs

Reason:

- Outliers/heavy tails for R treatment inflate the residual variance of R
- Leads to smaller scaling factor when robust methodology is applied
Summary (need for robust methodology):

- Degrees of freedom estimates for 2/8 studies (R treatment!) are small for both AUC and $C_{\text{max}}$.
- DIC prefers $\text{Bayes}_T$ to $\text{Bayes}_N$ in most cases.
- Bayesian CIs are generally somewhat wider than REML CIs.

Reason:

- Outliers/heavy tails for R treatment inflate the residual variance of R.
- Leads to smaller scaling factor when robust methodology is applied.
Summary (need for robust methodology):

- Degrees of freedom estimates for 2/8 studies (R treatment!) are small for both AUC and $C_{\text{max}}$
- DIC prefers $\text{Bayes}_T$ to $\text{Bayes}_N$ in most cases
- Bayesian CIs are generally somewhat wider than REML CIs

Reason:

- Outliers/heavy tails for R treatment inflate the residual variance of R
- Leads to smaller scaling factor when robust methodology is applied
Summary (need for robust methodology):

- Degrees of freedom estimates for 2/8 studies (R treatment!) are small for both AUC and $C_{\text{max}}$
- DIC prefers $\text{Bayes}_T$ to $\text{Bayes}_N$ in most cases
- Bayesian CIs are generally somewhat wider than REML CIs

Reason:

- Outliers/heavy tails for R treatment inflate the residual variance of R
- Leads to smaller scaling factor when robust methodology is applied
Summary (need for robust methodology):

- Degrees of freedom estimates for 2/8 studies (R treatment!) are small for both AUC and $C_{\text{max}}$
- DIC prefers $\text{Bayes}_T$ to $\text{Bayes}_N$ in most cases
- Bayesian CIs are generally somewhat wider than REML CIs

Reason:

- Outliers/heavy tails for R treatment inflate the residual variance of R
- Leads to smaller scaling factor when robust methodology is applied
### Table: AUC and $C_{\text{max}}$: RSABE Outcomes

<table>
<thead>
<tr>
<th>Method Outcome</th>
<th>REML</th>
<th>Bayes$_N$</th>
<th>Bayes$_T$</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pass</td>
<td>Fail</td>
<td>Pass</td>
<td>Fail</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>(%)</td>
<td>n</td>
<td>(%)</td>
</tr>
<tr>
<td>REML Pass</td>
<td>13 (86.7)</td>
<td>2 (13.3)</td>
<td>14 (93.3)</td>
<td>1 (6.7)</td>
</tr>
<tr>
<td>REML Fail</td>
<td>0 (0.0)</td>
<td>1 (100.0)</td>
<td>0 (0.0)</td>
<td>1 (100.0)</td>
</tr>
<tr>
<td>Bayes$_N$ Pass</td>
<td>13 (100.0)</td>
<td>0 (0.0)</td>
<td>13 (100.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>Bayes$_N$ Fail</td>
<td>2 (66.7)</td>
<td>1 (33.3)</td>
<td>1 (33.3)</td>
<td>2 (66.7)</td>
</tr>
<tr>
<td>Bayes$_T$ Pass</td>
<td>14 (100.0)</td>
<td>0 (0.0)</td>
<td>13 (92.9)</td>
<td>1 (7.1)</td>
</tr>
<tr>
<td>Bayes$_T$ Fail</td>
<td>1 (50.0)</td>
<td>1 (50.0)</td>
<td>0 (0.0)</td>
<td>2 (100.0)</td>
</tr>
<tr>
<td>Total</td>
<td>15 (93.8)</td>
<td>1 (6.3)</td>
<td>13 (81.3)</td>
<td>3 (18.8)</td>
</tr>
</tbody>
</table>

The DIC statistic preferred Bayes$_T$ over Bayes$_N$ in 14 out of 16 cases.
Summary (agreement between methods):

- Good agreement between REML and Bayes methods, in particular REML with Bayes-$\tau$