For $n = 27$ captured samples of the sirenian species *dugong* (sea cow), relate an animal’s length in meters, $Y_i$, to its age in years, $x_i$. 
For $n = 27$ captured samples of the sirenian species *dugong* (sea cow), relate an animal’s length in meters, $Y_i$, to its age in years, $x_i$.

To avoid a nonlinear model for now, transform $x_i$ to the log scale; plot of $Y$ versus $\log(x)$ looks fairly linear!
Simple linear regression in WinBUGS

\[ Y_i = \beta_0 + \beta_1 \log(x_i) + \epsilon_i, \quad i = 1, \ldots, n \]

where \( \epsilon_i \overset{iid}{\sim} N(0, \tau) \) and \( \tau = 1/\sigma^2 \), the precision in the data.

- Prior distributions:
  - flat for \( \beta_0, \beta_1 \)
  - vague gamma on \( \tau \) (say, \( \text{Gamma}(0.1, 0.1) \), which has mean 1 and variance 10) is traditional
Simple linear regression in WinBUGS

\[ Y_i = \beta_0 + \beta_1 \log(x_i) + \epsilon_i, \ i = 1, \ldots, n \]

where \( \epsilon_i \overset{iid}{\sim} N(0, \tau) \) and \( \tau = 1/\sigma^2 \), the precision in the data.

Prior distributions:
- flat for \( \beta_0, \beta_1 \)
- vague gamma on \( \tau \) (say, \( \text{Gamma}(0.1, 0.1) \), which has mean 1 and variance 10) is traditional

posterior correlation is reduced by centering the \( \log(x_i) \) around their own mean
Simple linear regression in WinBUGS

\[ Y_i = \beta_0 + \beta_1 \log(x_i) + \epsilon_i, \; i = 1, \ldots , n \]

where \( \epsilon_i \overset{iid}{\sim} N(0, \tau) \) and \( \tau = 1/\sigma^2 \), the precision in the data.

Prior distributions:

- flat for \( \beta_0, \beta_1 \)
- vague gamma on \( \tau \) (say, Gamma(0.1, 0.1), which has mean 1 and variance 10) is traditional

posterior correlation is reduced by centering the \( \log(x_i) \) around their own mean

Andrew Gelman suggests placing a uniform prior on \( \sigma \), bounding the prior away from 0 and \( \infty \implies U(0.01, 100) \)?
Simple linear regression in WinBUGS

\[ Y_i = \beta_0 + \beta_1 \log(x_i) + \epsilon_i, \ i = 1, \ldots, n \]

where \( \epsilon_i \overset{iid}{\sim} N(0, \tau) \) and \( \tau = 1/\sigma^2 \), the precision in the data.

- Prior distributions:
  - flat for \( \beta_0, \beta_1 \)
  - vague gamma on \( \tau \) (say, \( \text{Gamma}(0.1, 0.1) \), which has mean 1 and variance 10) is traditional
  - posterior correlation is reduced by centering the \( \log(x_i) \) around their own mean
  - Andrew Gelman suggests placing a \text{uniform} prior on \( \sigma \), bounding the prior away from 0 and \( \infty \Rightarrow U(.01, 100) \)?

- Code: 
  [www.biostat.umn.edu/~brad/data/dugongs_BUGS.txt](http://www.biostat.umn.edu/~brad/data/dugongs_BUGS.txt)
Model the untransformed dugong data as

\[ Y_i = \alpha - \beta \gamma^{x_i} + \epsilon_i, \ i = 1, \ldots, n, \]

where \( \alpha > 0, \beta > 0, 0 \leq \gamma \leq 1, \) and as usual \( \epsilon_i \overset{iid}{\sim} N(0, \tau) \) for \( \tau \equiv 1/\sigma^2 > 0. \)
Nonlinear regression in WinBUGS

In this model,
- $\alpha$ corresponds to the average length of a fully grown dugong ($x \to \infty$)
- $(\alpha - \beta)$ is the length of a dugong at birth ($x = 0$)
- $\gamma$ determines the growth rate: lower values produce an initially steep growth curve while higher values lead to gradual, almost linear growth.
Nonlinear regression in WinBUGS

In this model,

- $\alpha$ corresponds to the average length of a fully grown dugong ($x \to \infty$)
- $(\alpha - \beta)$ is the length of a dugong at birth ($x = 0$)
- $\gamma$ determines the growth rate: lower values produce an initially steep growth curve while higher values lead to gradual, almost linear growth.

Prior distributions: flat for $\alpha$ and $\beta$, $U(.01, 100)$ for $\sigma$, and $U(0.5, 1.0)$ for $\gamma$ (harder to estimate)
Nonlinear regression in WinBUGS

In this model,

- $\alpha$ corresponds to the average length of a fully grown dugong ($x \to \infty$)
- $(\alpha - \beta)$ is the length of a dugong at birth ($x = 0$)
- $\gamma$ determines the growth rate: lower values produce an initially steep growth curve while higher values lead to gradual, almost linear growth.

Prior distributions: flat for $\alpha$ and $\beta$, $U(0.01, 100)$ for $\sigma$, and $U(0.5, 1.0)$ for $\gamma$ (harder to estimate)

Code:
www.biostat.umn.edu/~brad/data/dugongsNL_BUGS.txt
Nonlinear regression in WinBUGS

In this model,

- $\alpha$ corresponds to the average length of a fully grown dugong ($x \to \infty$)
- $(\alpha - \beta)$ is the length of a dugong at birth ($x = 0$)
- $\gamma$ determines the growth rate: lower values produce an initially steep growth curve while higher values lead to gradual, almost linear growth.

Prior distributions: flat for $\alpha$ and $\beta$, $U(.01, 100)$ for $\sigma$, and $U(.5, 1.0)$ for $\gamma$ (harder to estimate)

Code:

www.biostat.umn.edu/~brad/data/dugongsNL_BUGS.txt

Obtain posterior density estimates and autocorrelation plots for $\alpha, \beta, \gamma,$ and $\sigma$, and investigate the bivariate posterior of $(\alpha, \gamma)$ using the Correlation tool on the Inference menu!
BUGS Example 3: Logistic Regression

Consider a binary version of the dugong data,

\[ Z_i = \begin{cases} 
1 & \text{if } Y_i > 2.4 \text{ (i.e., the dugong is “full-grown”) } \\
0 & \text{otherwise} 
\end{cases} \]
Consider a binary version of the dugong data,

\[ Z_i = \begin{cases} 
  1 & \text{if } Y_i > 2.4 \text{ (i.e., the dugong is “full-grown”)}, \\
  0 & \text{otherwise} 
\end{cases} \]

A logistic model for \( p_i = P(Z_i = 1) \) is then

\[ \logit(p_i) = \log\left[\frac{p_i}{1 - p_i}\right] = \beta_0 + \beta_1 \log(x_i) . \]
BUGS Example 3: Logistic Regression

Consider a binary version of the dugong data,

\[ Z_i = \begin{cases} 
1 & \text{if } Y_i > 2.4 \text{ (i.e., the dugong is “full-grown”) } \\
0 & \text{otherwise} 
\end{cases} \]

A logistic model for \( p_i = P(Z_i = 1) \) is then

\[
\text{logit}(p_i) = \log\left(\frac{p_i}{1 - p_i}\right) = \beta_0 + \beta_1 \log(x_i) .
\]

Two other commonly used link functions are the probit,

\[
\text{probit}(p_i) = \Phi^{-1}(p_i) = \beta_0 + \beta_1 \log(x_i) ,
\]

and the complementary log-log (cloglog),

\[
\text{cloglog}(p_i) = \log\left[-\log(1 - p_i)\right] = \beta_0 + \beta_1 \log(x_i) .
\]
Binary regression in WinBUGS

Code:
www.biostat.umn.edu/~brad/data/dugongsBin_BUGS.txt
Binary regression in WinBUGS

Code:
www.biostat.umn.edu/~brad/data/dugongsBin BUGS.txt

Code uses flat priors for $\beta_0$ and $\beta_1$, and the phi function, instead of the less stable probit function.
Binary regression in WinBUGS

- Code:
  www.biostat.umn.edu/~brad/data/dugongsBin_BUGS.txt

- Code uses flat priors for $\beta_0$ and $\beta_1$, and the phi function, instead of the less stable probit function.

- DIC scores for the three models:

<table>
<thead>
<tr>
<th>model</th>
<th>$\bar{D}$</th>
<th>$p_D$</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>logit</td>
<td>19.62</td>
<td>1.85</td>
<td>21.47</td>
</tr>
<tr>
<td>probit</td>
<td>19.30</td>
<td>1.87</td>
<td>21.17</td>
</tr>
<tr>
<td>cloglog</td>
<td>18.77</td>
<td>1.84</td>
<td>20.61</td>
</tr>
</tbody>
</table>

In fact, these scores can be obtained from a single run; see the “trick version” at the bottom of the BUGS file!
Binary regression in WinBUGS

- Code:
  www.biostat.umn.edu/~brad/data/dugongsBin_BUGS.txt

- Code uses flat priors for $\beta_0$ and $\beta_1$, and the phi function, instead of the less stable probit function.

- DIC scores for the three models:

<table>
<thead>
<tr>
<th>model</th>
<th>$\bar{D}$</th>
<th>$p_D$</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>logit</td>
<td>19.62</td>
<td>1.85</td>
<td>21.47</td>
</tr>
<tr>
<td>probit</td>
<td>19.30</td>
<td>1.87</td>
<td>21.17</td>
</tr>
<tr>
<td>cloglog</td>
<td>18.77</td>
<td>1.84</td>
<td>20.61</td>
</tr>
</tbody>
</table>

In fact, these scores can be obtained from a single run; see the “trick version” at the bottom of the BUGS file!

- Use the Comparison tool to compare the posteriors of $\beta_1$ across models, and the Correlation tool to check the bivariate posteriors of $(\beta_0, \beta_1)$ across models.
Fitted binary regression models

The logit and probit fits appear very similar, but the cloglog fitted curve is slightly different.
The logit and probit fits appear very similar, but the cloglog fitted curve is slightly different.

You can also compare $p_i$ posterior boxplots (induced by the link function and the $\beta_0$ and $\beta_1$ posteriors) using the Comparison tool.
Extend the usual two-stage (likelihood plus prior) Bayesian structure to a hierarchy of $L$ levels, where the joint distribution of the data and the parameters is

$$f(y|\theta_1)\pi_1(\theta_1|\theta_2)\pi_2(\theta_2|\theta_3)\cdots\pi_L(\theta_L|\lambda).$$
BUGS Example 4: Hierarchical Models

- Extend the usual two-stage (likelihood plus prior) Bayesian structure to a hierarchy of $L$ levels, where the joint distribution of the data and the parameters is

$$f(y|\theta_1)\pi_1(\theta_1|\theta_2)\pi_2(\theta_2|\theta_3)\cdots\pi_L(\theta_L|\lambda).$$

- $L$ is often determined by the number of subscripts on the data. For example, suppose $Y_{ijk}$ is the test score of child $k$ in classroom $j$ in school $i$ in a certain city. Model:

$$Y_{ijk}|\theta_{ij} \overset{\text{ind}}{\sim} N(\theta_{ij}, \tau_\theta) \quad (\theta_{ij} \text{ is the classroom effect})$$

$$\theta_{ij}|\eta_i \overset{\text{ind}}{\sim} N(\eta_i, \tau_\eta) \quad (\eta_i \text{ is the school effect})$$

$$\eta_i|\lambda \overset{\text{iid}}{\sim} N(\lambda, \tau_\lambda) \quad (\lambda \text{ is the grand mean})$$

Priors for $\lambda$ and the $\tau$’s now complete the specification!
Cross-Study (Meta-analysis) Data

Data: estimated log relative hazards $Y_{ij} = \hat{\beta}_{ij}$ obtained by fitting separate Cox proportional hazards regressions to the data from each of $J = 18$ clinical units participating in $I = 6$ different AIDS studies.
Cross-Study (Meta-analysis) Data

Data: estimated log relative hazards \( Y_{ij} = \beta_{ij} \) obtained by fitting separate Cox proportional hazards regressions to the data from each of \( J = 18 \) clinical units participating in \( I = 6 \) different AIDS studies.

To these data we wish to fit the cross-study model,

\[
Y_{ij} = a_i + b_j + s_{ij} + \epsilon_{ij}, \quad i = 1, \ldots, I, \quad j = 1, \ldots, J,
\]

where \( a_i = \) study main effect

\( b_j = \) unit main effect

\( s_{ij} = \) study-unit interaction term, and

\( \epsilon_{ij} \sim iid \ N(0, \sigma_{ij}^2) \)

and the estimated standard errors from the Cox regressions are used as (known) values of the \( \sigma_{ij} \).
# Cross-Study (Meta-analysis) Data

<table>
<thead>
<tr>
<th>Unit</th>
<th>Toxo</th>
<th>ddl/ddC</th>
<th>NuCombo ZDV+ddl</th>
<th>NuCombo ZDV+ddC</th>
<th>Fungal</th>
<th>CMV</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.814</td>
<td>NA</td>
<td>-0.406</td>
<td>0.298</td>
<td>0.094</td>
<td>NA</td>
</tr>
<tr>
<td>B</td>
<td>-0.203</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>C</td>
<td>-0.133</td>
<td>NA</td>
<td>0.218</td>
<td>-2.206</td>
<td>0.435</td>
<td>0.145</td>
</tr>
<tr>
<td>D</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>E</td>
<td>-0.715</td>
<td>-0.242</td>
<td>-0.544</td>
<td>-0.731</td>
<td>0.600</td>
<td>0.041</td>
</tr>
<tr>
<td>F</td>
<td>0.739</td>
<td>0.009</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>0.222</td>
</tr>
<tr>
<td>G</td>
<td>0.118</td>
<td>0.807</td>
<td>-0.047</td>
<td>0.913</td>
<td>-0.091</td>
<td>0.099</td>
</tr>
<tr>
<td>H</td>
<td>NA</td>
<td>-0.511</td>
<td>0.233</td>
<td>0.131</td>
<td>NA</td>
<td>0.017</td>
</tr>
<tr>
<td>I</td>
<td>NA</td>
<td>1.939</td>
<td>0.218</td>
<td>-0.066</td>
<td>NA</td>
<td>0.355</td>
</tr>
<tr>
<td>J</td>
<td>0.271</td>
<td>1.079</td>
<td>-0.277</td>
<td>-0.232</td>
<td>0.752</td>
<td>0.203</td>
</tr>
<tr>
<td>K</td>
<td>NA</td>
<td>NA</td>
<td>0.792</td>
<td>1.264</td>
<td>-0.357</td>
<td>0.807</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>1.217</td>
<td>0.165</td>
<td>0.385</td>
<td>0.172</td>
<td>-0.022</td>
<td>0.203</td>
</tr>
</tbody>
</table>
Cross-Study (Meta-analysis) Data

- Note that some values are missing ("NA") since
  - not all 18 units participated in all 6 studies
  - the Cox estimation procedure did not converge for some units that had few deaths
Cross-Study (Meta-analysis) Data

Note that some values are missing ("NA") since
not all 18 units participated in all 6 studies
the Cox estimation procedure did not converge for
some units that had few deaths

Goal: To identify which clinics are opinion leaders
(strongly agree with overall result across studies) and
which are dissenters (strongly disagree).
Cross-Study (Meta-analysis) Data

- Note that some values are missing ("NA") since
  - not all 18 units participated in all 6 studies
  - the Cox estimation procedure did not converge for some units that had few deaths

- **Goal:** To identify which clinics are opinion leaders (strongly agree with overall result across studies) and which are dissenters (strongly disagree).

  Here, overall results all favor the treatment (i.e. mostly negative $Y$'s) except in Trial 1 (Toxo). Thus we multiply all the $Y_{ij}$'s by $-1$ for $i \neq 1$, so that larger $Y_{ij}$ correspond in all cases to stronger agreement with the overall.
Cross-Study (Meta-analysis) Data

- Note that some values are missing ("NA") since not all 18 units participated in all 6 studies.
- The Cox estimation procedure did not converge for some units that had few deaths.

**Goal:** To identify which clinics are opinion leaders (strongly agree with overall result across studies) and which are dissenters (strongly disagree).

Here, overall results all favor the treatment (i.e. mostly negative $Y$’s) except in Trial 1 (Toxo). Thus we multiply all the $Y_{ij}$’s by $-1$ for $i \neq 1$, so that larger $Y_{ij}$ correspond in all cases to stronger agreement with the overall.

Next slide shows a plot of the $Y_{ij}$ values and associated approximate 95% CIs...
Cross-Study (Meta-analysis) Data

1: Toxo

2: ddl/ddC

3: NuCombo-ddl

4: NuCombo-ddC

5: Fungal

6: CMV
Cross-Study (Meta-analysis) Data

Second stage of our model:

\[ a_i \overset{iid}{\sim} N(0, 100^2), \quad b_j \overset{iid}{\sim} N(0, \sigma_b^2), \quad \text{and} \quad s_{ij} \overset{iid}{\sim} N(0, \sigma_s^2) \]
Cross-Study (Meta-analysis) Data

- Second stage of our model:

\[ a_i \overset{iid}{\sim} N(0, 100^2), \quad b_j \overset{iid}{\sim} N(0, \sigma_b^2), \quad \text{and} \quad s_{ij} \overset{iid}{\sim} N(0, \sigma_s^2) \]

- Third stage of our model:

\[ \sigma_b \sim Unif(0.01, 100) \quad \text{and} \quad \sigma_s \sim Unif(0.01, 100) \]

That is, we

- **preclude** borrowing of strength across studies, but
- **encourage** borrowing of strength across units
Cross-Study (Meta-analysis) Data

Second stage of our model:

\[ a_i \sim iid \ N(0, 100^2), \quad b_j \sim iid \ N(0, \sigma_b^2), \quad \text{and} \quad s_{ij} \sim iid \ N(0, \sigma_s^2) \]

Third stage of our model:

\[ \sigma_b \sim Unif(0.01, 100) \quad \text{and} \quad \sigma_s \sim Unif(0.01, 100) \]

That is, we

- preclude borrowing of strength across studies, but
- encourage borrowing of strength across units

With \( I + J + IJ \) parameters but fewer than \( IJ \) data points, some effects must be treated as random!
Cross-Study (Meta-analysis) Data

- Second stage of our model:

\[ a_i \overset{iid}{\sim} N(0, 100^2), \quad b_j \overset{iid}{\sim} N(0, \sigma_b^2), \quad \text{and} \quad s_{ij} \overset{iid}{\sim} N(0, \sigma_s^2) \]

- Third stage of our model:

\[ \sigma_b \sim Unif(0.01, 100) \quad \text{and} \quad \sigma_s \sim Unif(0.01, 100) \]

That is, we
- preclude borrowing of strength across studies, but
- encourage borrowing of strength across units

With \( I + J + IJ \) parameters but fewer than \( IJ \) data points, some effects must be treated as random!

- Code:
  www.biostat.umn.edu/~brad/data/crprot_BUGS.txt
Plot of $\theta_{ij}$ posterior means

◊ **Unit P** is an opinion leader; **Unit E** is a dissenter
Plot of $\theta_{ij}$ posterior means

◊ \textbf{Unit} \textit{P} is an opinion leader; \textbf{Unit} \textit{E} is a dissenter

◊ Substantial shrinkage towards 0 has occurred: mostly positive values; no estimated $\theta_{ij}$ greater than 0.6
Model Comparision via DIC

Since we lack replications for each study-unit \((i-j)\) combination, the interactions \(s_{ij}\) in this model were only weakly identified, and the model might well be better off without them (or even without the unit effects \(b_j\)).

As such, compare a variety of reduced models:

\[
Y[i,j] \sim \text{dnorm}(\theta[i,j], \Sigma[i,j])
\]

\[
\# \theta[i,j] \leftarrow a[i]+b[j]+s[i,j] \quad \# \text{full model}
\]

\[
\# \theta[i,j] \leftarrow a[i] + b[j] \quad \# \text{drop interactions}
\]

\[
\# \theta[i,j] \leftarrow a[i] + s[i,j] \quad \# \text{no unit effect}
\]

\[
\# \theta[i,j] \leftarrow b[j] + s[i,j] \quad \# \text{no study effect}
\]

\[
\# \theta[i,j] \leftarrow a[1] + b[j] \quad \# \text{unit + intercept}
\]

\[
\# \theta[i,j] \leftarrow b[j] \quad \# \text{unit effect only}
\]

\[
\theta[i,j] \leftarrow a[i] \quad \# \text{study effect only}
\]

Investigate \(p_D\) values for these models; are they consistent with posterior boxplots of the \(b_i\) and \(s_{ij}\)?
DIC results for Cross-Study Data:

<table>
<thead>
<tr>
<th>model</th>
<th>$\bar{D}$</th>
<th>$p_D$</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>full model</td>
<td>122.0</td>
<td>12.8</td>
<td>134.8</td>
</tr>
<tr>
<td>drop interactions</td>
<td>123.4</td>
<td>9.7</td>
<td>133.1</td>
</tr>
<tr>
<td>no unit effect</td>
<td>123.8</td>
<td>10.0</td>
<td>133.8</td>
</tr>
<tr>
<td>no study effect</td>
<td>121.4</td>
<td>9.7</td>
<td>131.1</td>
</tr>
<tr>
<td>unit + intercept</td>
<td>120.3</td>
<td>4.6</td>
<td>124.9</td>
</tr>
<tr>
<td>unit effect only</td>
<td>122.9</td>
<td>6.2</td>
<td>129.1</td>
</tr>
<tr>
<td>study effect only</td>
<td>126.0</td>
<td>6.0</td>
<td>132.0</td>
</tr>
</tbody>
</table>

The **DIC-best model** is the one with only an intercept (a role played here by $a_1$) and the unit effects $b_j$.

These DIC differences are not much larger than their possible Monte Carlo errors, so almost any of these models could be justified here.
BUGS Example 5: Survival Modeling

Our data arises from a clinical trial comparing two treatments for *Mycobacterium avium complex (MAC)*, a disease common in late stage HIV-infected persons. Eleven clinical centers ("units") have enrolled a total of 69 patients in the trial, of which 18 have died.
BUGS Example 5: Survival Modeling

Our data arises from a clinical trial comparing two treatments for *Mycobacterium avium* complex (MAC), a disease common in late stage HIV-infected persons. Eleven clinical centers (“units”) have enrolled a total of 69 patients in the trial, of which 18 have died.

For \( j = 1, \ldots, n_i \) and \( i = 1, \ldots, k \), let

\[
\begin{align*}
t_{ij} & \quad \text{time to death or censoring} \\
x_{ij} & \quad \text{treatment indicator for subject j in stratum i}
\end{align*}
\]
BUGS Example 5: Survival Modeling

Our data arises from a clinical trial comparing two treatments for *Mycobacterium avium complex (MAC)*, a disease common in late stage HIV-infected persons. Eleven clinical centers ("units") have enrolled a total of 69 patients in the trial, of which 18 have died.

For $j = 1, \ldots, n_i$ and $i = 1, \ldots, k$, let

\[
\begin{align*}
t_{ij} &= \text{time to death or censoring} \\
x_{ij} &= \text{treatment indicator for subject } j \text{ in stratum } i
\end{align*}
\]

Next page gives survival times (in half-days) from the MAC treatment trial, where "+" indicates a censored observation...
### MAC Survival Data

<table>
<thead>
<tr>
<th>unit</th>
<th>drug</th>
<th>time</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>74+</td>
</tr>
<tr>
<td>A</td>
<td>2</td>
<td>248</td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>272+</td>
</tr>
<tr>
<td>A</td>
<td>2</td>
<td>344</td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>4+</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>156+</td>
</tr>
<tr>
<td>C</td>
<td>2</td>
<td>100+</td>
</tr>
<tr>
<td>D</td>
<td>2</td>
<td>20+</td>
</tr>
<tr>
<td>D</td>
<td>2</td>
<td>64</td>
</tr>
<tr>
<td>D</td>
<td>2</td>
<td>88</td>
</tr>
<tr>
<td>D</td>
<td>2</td>
<td>148+</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>unit</th>
<th>drug</th>
<th>time</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>1</td>
<td>214</td>
</tr>
<tr>
<td>E</td>
<td>2</td>
<td>228+</td>
</tr>
<tr>
<td>E</td>
<td>2</td>
<td>262</td>
</tr>
<tr>
<td>F</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>F</td>
<td>2</td>
<td>16+</td>
</tr>
<tr>
<td>F</td>
<td>1</td>
<td>76</td>
</tr>
<tr>
<td>F</td>
<td>2</td>
<td>80</td>
</tr>
<tr>
<td>F</td>
<td>2</td>
<td>202</td>
</tr>
<tr>
<td>F</td>
<td>1</td>
<td>258+</td>
</tr>
<tr>
<td>F</td>
<td>1</td>
<td>268+</td>
</tr>
<tr>
<td>I</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>I</td>
<td>2</td>
<td>16+</td>
</tr>
<tr>
<td>I</td>
<td>2</td>
<td>40</td>
</tr>
<tr>
<td>I</td>
<td>1</td>
<td>120+</td>
</tr>
<tr>
<td>I</td>
<td>1</td>
<td>168+</td>
</tr>
<tr>
<td>I</td>
<td>2</td>
<td>174+</td>
</tr>
<tr>
<td>I</td>
<td>1</td>
<td>268+</td>
</tr>
<tr>
<td>I</td>
<td>2</td>
<td>276</td>
</tr>
<tr>
<td>I</td>
<td>1</td>
<td>286+</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>K</td>
<td>2</td>
<td>106+</td>
</tr>
</tbody>
</table>
MAC Survival Data

With proportional hazards and a Weibull baseline hazard, stratum \(i\)'s hazard is

\[
h(t_{ij}; x_{ij}) = h_0(t_{ij}) \omega_i \exp(\beta_0 + \beta_1 x_{ij}) \\
= \rho_i t_{ij}^{\rho_i - 1} \exp(\beta_0 + \beta_1 x_{ij} + W_i),
\]

where \(\rho_i > 0\), \(\beta = (\beta_0, \beta_1)' \in \mathbb{R}^2\), and \(W_i = \log \omega_i\) is a clinic-specific frailty term.
MAC Survival Data

With proportional hazards and a Weibull baseline hazard, stratum $i$’s hazard is

$$h(t_{ij}; x_{ij}) = h_0(t_{ij})\omega_i \exp(\beta_0 + \beta_1 x_{ij})$$

$$= \rho_i t_{ij}^{\rho_i-1} \exp(\beta_0 + \beta_1 x_{ij} + W_i),$$

where $\rho_i > 0$, $\beta = (\beta_0, \beta_1)' \in \mathbb{R}^2$, and $W_i = \log \omega_i$ is a clinic-specific frailty term.

The $W_i$ capture overall differences among the clinics, while the $\rho_i$ allow differing baseline hazards which either increase ($\rho_i > 1$) or decrease ($\rho_i < 1$) over time. We assume i.i.d. specifications for these random effects,

$$W_i \overset{iid}{\sim} N(0, 1/\tau) \quad \text{and} \quad \rho_i \overset{iid}{\sim} G(\alpha, \alpha).$$
MAC Survival Data

As in the \textit{mice} example (\textsc{WinBUGS} Examples Vol 1),

\[ \mu_{ij} = \exp(\beta_0 + \beta_1 x_{ij} + W_i), \]

so that

\[ t_{ij} \sim \text{Weibull}(\rho_i, \mu_{ij}). \]
MAC Survival Data

As in the mice example (WinBUGS Examples Vol 1),

$$\mu_{ij} = \exp(\beta_0 + \beta_1 x_{ij} + W_i),$$

so that

$$t_{ij} \sim \text{Weibull}(\rho_i, \mu_{ij}).$$

We recode the drug covariate from (1,2) to (−1,1) (i.e., set $x_{ij} = 2\text{drug}_{ij} - 3$) to ease collinearity between the slope $\beta_1$ and the intercept $\beta_0$. 
MAC Survival Data

- As in the mice example (WinBUGS Examples Vol 1),
  \[ \mu_{ij} = \exp(\beta_0 + \beta_1 x_{ij} + W_i), \]
  so that
  \[ t_{ij} \sim \text{Weibull}(\rho_i, \mu_{ij}). \]
- We recode the drug covariate from (1,2) to (–1,1) (i.e., set \( x_{ij} = 2 \text{drug}_{ij} - 3 \)) to ease collinearity between the slope \( \beta_1 \) and the intercept \( \beta_0 \).
- We place vague priors on \( \beta_0 \) and \( \beta_1 \), a moderately informative \( G(1, 1) \) prior on \( \tau \), and set \( \alpha = 10 \).
MAC Survival Data

As in the mice example (WinBUGS Examples Vol 1),

\[ \mu_{ij} = \exp(\beta_0 + \beta_1 x_{ij} + W_i), \]

so that

\[ t_{ij} \sim \text{Weibull}(\rho_i, \mu_{ij}). \]

We recode the drug covariate from (1,2) to (–1,1) (i.e., set \( x_{ij} = 2\text{drug}_{ij} - 3 \)) to ease collinearity between the slope \( \beta_1 \) and the intercept \( \beta_0 \).

We place vague priors on \( \beta_0 \) and \( \beta_1 \), a moderately informative \( G(1, 1) \) prior on \( \tau \), and set \( \alpha = 10 \).

Data: www.biostat.umn.edu/~brad/data/MAC.dat
Code: www.biostat.umn.edu/~brad/data/MACfrailty_BUGS.txt
## MAC Survival Results

<table>
<thead>
<tr>
<th>node (unit)</th>
<th>mean</th>
<th>sd</th>
<th>MC error</th>
<th>2.5%</th>
<th>median</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W_1$ (A)</td>
<td>-0.04912</td>
<td>0.835</td>
<td>0.02103</td>
<td>-1.775</td>
<td>-0.04596</td>
<td>1.639</td>
</tr>
<tr>
<td>$W_3$ (C)</td>
<td>-0.1829</td>
<td>0.9173</td>
<td>0.01782</td>
<td>-2.2</td>
<td>-0.1358</td>
<td>1.52</td>
</tr>
<tr>
<td>$W_5$ (E)</td>
<td>-0.03198</td>
<td>0.8107</td>
<td>0.03193</td>
<td>-1.682</td>
<td>-0.02653</td>
<td>1.572</td>
</tr>
<tr>
<td>$W_6$ (F)</td>
<td>0.4173</td>
<td>0.8277</td>
<td>0.04065</td>
<td>-1.066</td>
<td>0.3593</td>
<td>2.227</td>
</tr>
<tr>
<td>$W_9$ (I)</td>
<td>0.2546</td>
<td>0.7969</td>
<td>0.03694</td>
<td>-1.241</td>
<td>0.2164</td>
<td>1.968</td>
</tr>
<tr>
<td>$W_{11}$ (K)</td>
<td>-0.1945</td>
<td>0.9093</td>
<td>0.02093</td>
<td>-2.139</td>
<td>-0.1638</td>
<td>1.502</td>
</tr>
</tbody>
</table>

| $\rho_1$ (A) | 1.086 | 0.1922 | 0.007168 | 0.7044 | 1.083   | 1.474  |
| $\rho_3$ (C) | 0.9008 | 0.2487 | 0.006311 | 0.4663 | 0.8824  | 1.431  |
| $\rho_5$ (E) | 1.143 | 0.1887 | 0.00958 | 0.7904 | 1.139   | 1.521  |
| $\rho_6$ (F) | 0.935 | 0.1597 | 0.008364 | 0.6321 | 0.931   | 1.265  |
| $\rho_9$ (I) | 0.9788 | 0.1683 | 0.008735 | 0.6652 | 0.9705  | 1.339  |
| $\rho_{11}$ (K) | 0.8807 | 0.2392 | 0.01034 | 0.4558 | 0.8612  | 1.394  |

| $\tau$ | 1.733 | 1.181 | 0.03723 | 0.3042 | 1.468   | 4.819  |
| $\beta_0$ | -7.111 | 0.689 | 0.04474 | -8.552 | -7.073  | -5.874 |
| $\beta_1$ | 0.596 | 0.2964 | 0.01048 | 0.06099 | 0.5783  | 1.245  |
| $RR$ | 3.98 | 2.951 | 0.1122 | 1.13   | 3.179   | 12.05  |
MAC Survival Results

 Units A and E have moderate overall risk ($W_i ≈ 0$) but increasing hazards ($\rho > 1$): few deaths, but they occur late
MAC Survival Results

- Units A and E have moderate overall risk \((W_i \approx 0)\) but increasing hazards \((\rho > 1)\): few deaths, but they occur late

- Units F and I have high overall risk \((W_i > 0)\) but decreasing hazards \((\rho < 1)\): several early deaths, many long-term survivors
MAC Survival Results

- Units A and E have **moderate overall risk** \( W_i \approx 0 \) but **increasing hazards** \( \rho > 1 \): few deaths, but they occur late

- Units F and I have **high overall risk** \( W_i > 0 \) but **decreasing hazards** \( \rho < 1 \): several early deaths, many long-term survivors

- Units C and K have **low overall risk** \( W_i < 0 \) and **decreasing hazards** \( \rho < 1 \): no deaths at all; a few survivors
MAC Survival Results

- Units A and E have moderate overall risk \((W_i \approx 0)\) but increasing hazards \((\rho > 1)\): few deaths, but they occur late

- Units F and I have high overall risk \((W_i > 0)\) but decreasing hazards \((\rho < 1)\): several early deaths, many long-term survivors

- Units C and K have low overall risk \((W_i < 0)\) and decreasing hazards \((\rho < 1)\): no deaths at all; a few survivors

- Drugs differ significantly: CI for \(\beta_1 (RR)\) excludes 0 (1)
MAC Survival Results

- Units A and E have moderate overall risk \((W_i \approx 0)\) but increasing hazards \((\rho > 1)\): few deaths, but they occur late

- Units F and I have high overall risk \((W_i > 0)\) but decreasing hazards \((\rho < 1)\): several early deaths, many long-term survivors

- Units C and K have low overall risk \((W_i < 0)\) and decreasing hazards \((\rho < 1)\): no deaths at all; a few survivors

- Drugs differ significantly: CI for \(\beta_1 \ (RR)\) excludes 0 (1)

Note: This has all been for two sets of random effects \((W_i \text{ and } \rho_i)\), called “Model 2” in the BUGS code. You will also see models having three (adding \(\beta_{1i}\), one (deleting \(\rho_i\)), or zero sets of random effects!
BRugs Example 1: Model assessment

Basic tool here is the cross-validation residual

\[ r_i = y_i - E(y_i | y(i)) \]

where \( y(i) \) denotes the vector of all the data except the \( i^{th} \) value, i.e.

\[ y(i) = (y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_n)' \]

Outliers are indicated by large standardized residuals,

\[ d_i = r_i / \sqrt{Var(y_i | y(i))} \].
BRugs Example 1: Model assessment

Basic tool here is the cross-validation residual

\[ r_i = y_i - E(y_i | y(i)) , \]

where \( y(i) \) denotes the vector of all the data except the \( i^{th} \) value, i.e.

\[ y(i) = (y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_n)' . \]

Outliers are indicated by large standardized residuals,

\[ d_i = r_i / \sqrt{Var(y_i | y(i))} . \]

Also of interest is the conditional predictive ordinate,

\[ p(y_i | y(i)) = \int p(y_i | \theta, y(i)) p(\theta | y(i)) d\theta , \]

the height of the conditional density at the observed value of \( y_i \).

\[ \rightarrow \] large values indicate good prediction of \( y_i \).
Residuals: Approximate method

Using MC draws $\theta^{(g)} \sim p(\theta|y)$, we have

$$E(y_i|y_{(i)}) = \int \int y_i f(y_i|\theta)p(\theta|y_{(i)}) dy_i d\theta$$

$$= \int E(y_i|\theta)p(\theta|y_{(i)}) d\theta$$

$$\approx \int E(y_i|\theta)p(\theta|y) d\theta$$

$$\approx \frac{1}{G} \sum_{g=1}^{G} E(y_i|\theta^{(g)})$$.
Residuals: Approximate method

Using MC draws $\theta^{(g)} \sim p(\theta|y)$, we have

$$E(y_i|y_{(i)}) = \int \int y_i f(y_i|\theta)p(\theta|y_{(i)}) dy_i d\theta$$

$$= \int E(y_i|\theta)p(\theta|y_{(i)}) d\theta$$

$$\approx \int E(y_i|\theta)p(\theta|y) d\theta$$

$$\approx \frac{1}{G} \sum_{g=1}^{G} E(y_i|\theta^{(g)}) .$$

Approximation should be adequate unless the dataset is small and $y_i$ is an extreme outlier.
Residuals: Approximate method

Using MC draws $\theta^{(g)} \sim p(\theta|y)$, we have

$$E(y_i|y_{(i)}) = \int \int y_i f(y_i|\theta)p(\theta|y_{(i)}) dy_i d\theta$$

$$= \int E(y_i|\theta)p(\theta|y_{(i)}) d\theta$$

$$\approx \int E(y_i|\theta)p(\theta|y) d\theta$$

$$\approx \frac{1}{G} \sum_{g=1}^{G} E(y_i|\theta^{(g)})$$.

Approximation should be adequate unless the dataset is small and $y_i$ is an extreme outlier.

Same $\theta^{(g)}$’s may be used for each $i = 1, \ldots, n$. 
The ratio to compute the standardized residuals $d_i$ must be done outside of WinBUGS. Might instead define

$$d_i^* = \frac{y_i - E(y_i|\theta)}{\sqrt{\text{Var}(y_i|\theta)}}.$$

We then find $E(d_i^*|y)$, the posterior average of the ratio (instead of the ratio of the posterior averages).
Approximate methods in WinBUGS

- The ratio to compute the standardized residuals $d_i$ must be done outside of WinBUGS. Might instead define

$$d_i^* = \frac{y_i - E(y_i \mid \theta)}{\sqrt{Var(y_i \mid \theta)}}.$$ 

We then find $E(d_i^* \mid y)$, the posterior average of the ratio (instead of the ratio of the posterior averages).

- For the exact method, we must evaluate $E(y_i \mid y(i))$ and $Var(y_i \mid y(i))$ separately. For the latter, use the facts that

$$Var(y_i \mid y(i)) = E(y_i^2 \mid y(i)) - [E(y_i \mid y(i))]^2,$n

and

$$E(y_i^2 \mid y(i)) = \int E(y_i^2 \mid \theta)p(\theta \mid y(i))d\theta$$

$$= \int \{Var(y_i \mid \theta) + [E(y_i \mid \theta)]^2\}p(\theta \mid y(i))d\theta.$$
Residuals: Exact method

An exact solution then arises by calling \texttt{WinBUGS} \( n \) times, once for each “leave one out” dataset!
Residuals: Exact method

- An exact solution then arises by calling WinBUGS $n$ times, once for each “leave one out” dataset!
- This can be accomplished in BRugs, a suite of R routines for calling OpenBUGS from R, originally written by WinBUGS head programmer Andrew Thomas, and refined and maintained by Uwe Ligges
Residuals: Exact method

- An exact solution then arises by calling WinBUGS $n$ times, once for each “leave one out” dataset!

- This can be accomplished in BRugs, a suite of R routines for calling OpenBUGS from R, originally written by WinBUGS head programmer Andrew Thomas, and refined and maintained by Uwe Ligges

- All necessary programs and instructions can be downloaded from www.biostat.umn.edu/~brad/software/BRugs
Residuals: Exact method

- An exact solution then arises by calling WinBUGS \( n \) times, once for each “leave one out” dataset!
- This can be accomplished in BRugs, a suite of R routines for calling OpenBUGS from R, originally written by WinBUGS head programmer Andrew Thomas, and refined and maintained by Uwe Ligges
- All necessary programs and instructions can be downloaded from www.biostat.umn.edu/~brad/software/BRugs
- Note that we will now have both:
  - an R program that organizes the dataset, contains all the BRugs commands, and summarizes the output
  - a piece of BUGS code that is sent by R to OpenBUGS
Numerical illustration: Stack Loss data

An oft-analyzed dataset, featuring the stack loss \( Y \) (ammonia escaping), and three covariates \( X_1 \) (air flow), \( X_2 \) (temperature), and \( X_3 \) (acid concentration).
Numerical illustration: Stack Loss data

- An oft-analyzed dataset, featuring the stack loss $Y$ (ammonia escaping), and three covariates $X_1$ (air flow), $X_2$ (temperature), and $X_3$ (acid concentration).

- Fit the linear regression model

$$ Y_i \sim N(\beta_0 + \beta_1 z_{i1} + \beta_2 z_{i2} + \beta_3 z_{i3} , \tau), $$

where the $z_{ij}$ are the standardized covariates. We take flat priors on the $\beta$s and a Gelman-style noninformative prior on $\sigma = 1/\sqrt{\tau}$. 
Numerical illustration: Stack Loss data

- An oft-analyzed dataset, featuring the stack loss $Y$ (ammonia escaping), and three covariates $X_1$ (air flow), $X_2$ (temperature), and $X_3$ (acid concentration).

- Fit the linear regression model

$$Y_i \sim N(\beta_0 + \beta_1 z_{i1} + \beta_2 z_{i2} + \beta_3 z_{i3}, \tau),$$

where the $z_{ij}$ are the standardized covariates. We take flat priors on the $\beta$s and a Gelman-style noninformative prior on $\sigma = 1/\sqrt{\tau}$.

- WinBUGS code and data for approximate method: www.biostat.umn.edu/~brad/data/stacks_BUGS.txt

- BRugs code and data for exact method: www.biostat.umn.edu/~brad/software/BRugs
Numerical illustration: Stack Loss data

- An oft-analyzed dataset, featuring the stack loss $Y$ (ammonia escaping), and three covariates $X_1$ (air flow), $X_2$ (temperature), and $X_3$ (acid concentration).

- Fit the linear regression model

$$Y_i \sim N\left(\beta_0 + \beta_1 z_{i1} + \beta_2 z_{i2} + \beta_3 z_{i3}, \tau\right),$$

where the $z_{ij}$ are the standardized covariates. We take flat priors on the $\beta$s and a Gelman-style noninformative prior on $\sigma = 1/\sqrt{\tau}$.

- **WinBUGS** code and data for approximate method:
  
  www.biostat.umn.edu/~brad/data/stacks_BUGS.txt

- **BRugs** code and data for exact method:
  
  www.biostat.umn.edu/~brad/software/BRugs

- See also “stacks” in **WinBUGS Examples Volume I**!
Approximate vs. Exact Results

<table>
<thead>
<tr>
<th>obs</th>
<th>sresid approx</th>
<th>sresid exact</th>
<th>CPO approx</th>
<th>CPO exact</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.948</td>
<td>1.098</td>
<td>0.178</td>
<td>0.124</td>
</tr>
<tr>
<td>2</td>
<td>−0.566</td>
<td>−0.628</td>
<td>0.224</td>
<td>0.188</td>
</tr>
<tr>
<td>3</td>
<td>1.337</td>
<td>1.461</td>
<td>0.122</td>
<td>0.084</td>
</tr>
<tr>
<td>4</td>
<td>1.672</td>
<td>1.851</td>
<td>0.078</td>
<td>0.047</td>
</tr>
<tr>
<td>5</td>
<td>−0.504</td>
<td>−0.477</td>
<td>0.251</td>
<td>0.244</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>21</td>
<td>−2.126</td>
<td>−3.012</td>
<td>0.046</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Approximate residuals are too small, especially for the most outlying observations!
## Approximate vs. Exact Results

<table>
<thead>
<tr>
<th>obs</th>
<th>sresid approx</th>
<th>sresid exact</th>
<th>CPO approx</th>
<th>CPO exact</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.948</td>
<td>1.098</td>
<td>0.178</td>
<td>0.124</td>
</tr>
<tr>
<td>2</td>
<td>-0.566</td>
<td>-0.628</td>
<td>0.224</td>
<td>0.188</td>
</tr>
<tr>
<td>3</td>
<td>1.337</td>
<td>1.461</td>
<td>0.122</td>
<td>0.084</td>
</tr>
<tr>
<td>4</td>
<td>1.672</td>
<td>1.851</td>
<td>0.078</td>
<td>0.047</td>
</tr>
<tr>
<td>5</td>
<td>-0.504</td>
<td>-0.477</td>
<td>0.251</td>
<td>0.244</td>
</tr>
<tr>
<td></td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>21</td>
<td>-2.126</td>
<td>-3.012</td>
<td>0.046</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Approximate residuals are too small, especially for the most outlying observations!

Approximate CPOs also tend to understate lack of fit
Following our MAC survival model, let $t_i$ be the time until death for subject $i$, with corresponding treatment indicator $x_i$ ($= 0$ or $1$ for control and treatment, respectively). Suppose

$$t_i \sim \text{Weibull}(r, \mu_i), \text{ where } \mu_i = e^{-(\beta_0 + \beta_1 x_i)}.$$
Following our MAC survival model, let $t_i$ be the time until death for subject $i$, with corresponding treatment indicator $x_i$ (= 0 or 1 for control and treatment, respectively). Suppose

$$t_i \sim \text{Weibull}(r, \mu_i), \quad \text{where} \quad \mu_i = e^{-(\beta_0 + \beta_1 x_i)}.$$ 

Then the baseline hazard function is $\lambda_0(t_i) = rt_i^{r-1}$, and the median survival time for subject $i$ is

$$m_i = \left[(\log 2)e^{\beta_0 + \beta_1 x_i}\right]^{1/r}.$$
Following our MAC survival model, let $t_i$ be the time until death for subject $i$, with corresponding treatment indicator $x_i$ (= 0 or 1 for control and treatment, respectively). Suppose

$$t_i \sim \text{Weibull}(r, \mu_i), \text{ where } \mu_i = e^{-\left(\beta_0 + \beta_1 x_i\right)}.$$ 

Then the baseline hazard function is $\lambda_0(t_i) = rt_i^{r-1}$, and the median survival time for subject $i$ is

$$m_i = \left[(\log 2)e^{\beta_0 + \beta_1 x_i}\right]^{1/r}.$$ 

The value of $\beta_1$ corresponding to a 15% increase in median survival in the treatment group satisfies

$$e^{\beta_1/r} = 1.15 \iff \beta_1 = r \log(1.15).$$
Range of equivalence

- The range of $\beta_1$ values within which we are indifferent as to use of treatment or control.
Range of equivalence

- The range of $\beta_1$ values within which we are indifferent as to use of treatment or control
- lower limit $\beta_I$, the clinical inferiority boundary
  - We typically take $\beta_I = 0$, since we would never prefer a harmful treatment
Range of equivalence

- The range of $\beta_1$ values within which we are indifferent as to use of treatment or control

- lower limit $\beta_I$, the **clinical inferiority** boundary
  - We typically take $\beta_I = 0$, since we would never prefer a harmful treatment

- upper limit $\beta_S$, the **clinical superiority** boundary
  - We typically take $\beta_S > 0$, since we may require “clinically significant” improvement under the treatment (due to cost, toxicity, etc.)
  
  **Example:** If $r = 2$, then $\beta_S = 2 \log(1.15) \approx 0.28$ corresponds to 15% improvement in median survival
Range of equivalence

- The range of $\beta_1$ values within which we are indifferent as to use of treatment or control
- lower limit $\beta_I$, the clinical inferiority boundary
  - We typically take $\beta_I = 0$, since we would never prefer a harmful treatment
- upper limit $\beta_S$, the clinical superiority boundary
  - We typically take $\beta_S > 0$, since we may require "clinically significant" improvement under the treatment (due to cost, toxicity, etc.)
  - Example: If $r = 2$, then $\beta_S = 2 \log(1.15) \approx 0.28$ corresponds to 15% improvement in median survival
- The outcome of the trial can then be based on the location of the 95% posterior confidence interval for $\beta_1$, say $(\beta_L, \beta_U)$, relative to the indifference zone!....
The six possible outcomes and decisions

- Accept control \( (\beta_L, \beta_U) \)
- Reject treatment \( (\beta_L, \beta_U) \)
- Equivalence \( (\beta_L, \beta_U) \)
- Reject control \( (\beta_L, \beta_U) \)
- Accept treatment \( (\beta_L, \beta_U) \)
- No decision

Note both “acceptance” and “rejection” are possible!

- \( \beta_I = 0 \)
- \( \beta_S = 0.28 \)
Community of priors

Spiegelhalter et al. (1994) recommend considering several priors, in order to represent the broadest possible audience:

- **Skeptical Prior**
  
  One that believes the treatment is likely no better than control (as might be believed by the FDA)
Community of priors

Spiegelhalter et al. (1994) recommend considering several priors, in order to represent the broadest possible audience:

- **Skeptical** Prior
  - One that believes the treatment is likely no better than control (as might be believed by the FDA)

- **Enthusiastic (or Clinical)** Prior
  - One that believes the treatment will succeed (typical of the clinicians running the trial)
Community of priors

Spiegelhalter et al. (1994) recommend considering several priors, in order to represent the broadest possible audience:

- **Skeptical Prior**
  - One that believes the treatment is likely no better than control (as might be believed by the FDA)

- **Enthusiastic (or Clinical) Prior**
  - One that believes the treatment will succeed (typical of the clinicians running the trial)

- **Reference (or Noninformative) Prior**
  - One that expresses no particular opinion about the treatment’s merit
  - Often a improper uniform (“flat”) prior is permissible
MCMC-based Bayesian design

Simulating the power or other operating characteristics (say, Type I error) in this setting works as follows:

- Sample “true” $\beta$ values from an assumed “true prior” (skeptical, enthusiastic, or in between)
MCMC-based Bayesian design

Simulating the power or other operating characteristics (say, Type I error) in this setting works as follows:

- Sample “true” $\beta$ values from an assumed “true prior” (skeptical, enthusiastic, or in between)

- Given these, sample fake survival times $t_i$ (say, $N$ from each study group) from the Weibull
Simulating the power or other operating characteristics (say, Type I error) in this setting works as follows:

- Sample “true” \( \beta \) values from an assumed “true prior” (skeptical, enthusiastic, or in between)
- Given these, sample fake survival times \( t_i \) (say, \( N \) from each study group) from the Weibull
- We may also wish to sample fake censoring times \( c_i \) from a particular distribution (e.g., a normal truncated below 0); for all \( i \) such that \( t_i > c_i \), replace \( t_i \) by “NA”
MCMC-based Bayesian design

Simulating the power or other operating characteristics (say, Type I error) in this setting works as follows:

- Sample “true” $\beta$ values from an assumed “true prior” (skeptical, enthusiastic, or in between)
- Given these, sample fake survival times $t_i$ (say, $N$ from each study group) from the Weibull
- We may also wish to sample fake censoring times $c_i$ from a particular distribution (e.g., a normal truncated below 0); for all $i$ such that $t_i > c_i$, replace $t_i$ by “NA”
- Compute $(\beta_L, \beta_U)$ by calling OpenBUGS from R
MCMC-based Bayesian design

Simulating the power or other operating characteristics (say, Type I error) in this setting works as follows:

- Sample “true” $\beta$ values from an assumed “true prior” (skeptical, enthusiastic, or in between)
- Given these, sample fake survival times $t_i$ (say, $N$ from each study group) from the Weibull
- We may also wish to sample fake censoring times $c_i$ from a particular distribution (e.g., a normal truncated below 0); for all $i$ such that $t_i > c_i$, replace $t_i$ by “NA”
- Compute $(\beta_L, \beta_U)$ by calling OpenBUGS from R
- Determine the simulated trial’s outcome based on location of $(\beta_L, \beta_U)$ relative to the indifference zone
MCMC-based Bayesian design

Simulating the power or other operating characteristics (say, Type I error) in this setting works as follows:

- Sample “true” $\beta$ values from an assumed “true prior” (skeptical, enthusiastic, or in between)

- Given these, sample fake survival times $t_i$ (say, $N$ from each study group) from the Weibull

- We may also wish to sample fake censoring times $c_i$ from a particular distribution (e.g., a normal truncated below 0); for all $i$ such that $t_i > c_i$, replace $t_i$ by “NA”

- Compute $(\beta_L, \beta_U)$ by calling OpenBUGS from R

- Determine the simulated trial’s outcome based on location of $(\beta_L, \beta_U)$ relative to the indifference zone

- Repeat this process $N_{rep}$ times; report empirical frequencies of the six possible outcomes
Results from Power.BRugs

Assuming:

- Weibull shape $r = 2$, and $N = 50$ in each group
- median survival of 36 days with 50% improvement in the treatment group
- a $N(80, 20)$ censoring distribution
- the enthusiastic prior as the “truth”

We obtain the following output from $N_{rep} = 100$ reps:
Results from Power.BRugs

Assuming:
- Weibull shape $r = 2$, and $N = 50$ in each group
- median survival of 36 days with 50% improvement in the treatment group
- a $N(80, 20)$ censoring distribution
- the enthusiastic prior as the “truth”

We obtain the following output from $N_{rep} = 100$ reps:

Here are simulated outcome frequencies for $N= 50$
- accept control: 0
- reject treatment: 0.07
- equivalence: 0
- reject control: 0.87
- accept treatment: 0.06
- no decision: 0

End of BRugs power simulation
Homework Problems

WinBUGS

PK hierarchical linear model:
www.biostat.umn.edu/~brad/data/PK_BUGS.txt

PK hierarchical nonlinear model:
www.biostat.umn.edu/~brad/data/PKNL_BUGS.txt

Interstim multivariate model:
www.biostat.umn.edu/~brad/data/InterStim.odc

Bayesian $p$-values (illustrated with stacks data):
www.biostat.umn.edu/~brad/data/stackspval_BUGS.txt
Homework Problems

- **WinBUGS**
  - PK hierarchical linear model:
    www.biostat.umn.edu/~brad/data/PK_BUGS.txt
  - PK hierarchical nonlinear model:
    www.biostat.umn.edu/~brad/data/PKNL_BUGS.txt
  - Interstim multivariate model:
    www.biostat.umn.edu/~brad/data/InterStim.odc
  - Bayesian $p$-values (illustrated with stacks data):
    www.biostat.umn.edu/~brad/data/stackspval_BUGS.txt

- **BRugs**
  - Design for binary and Cox PH models: Brian Hobbs’ webpage:
    www.biostat.umn.edu/~brianho/papers/2007/JBS/prac_bayes_design.html
Homework Problems

**WinBUGS**
- PK hierarchical linear model:  
  www.biostat.umn.edu/~brad/data/PK_BUGS.txt
- PK hierarchical nonlinear model:  
  www.biostat.umn.edu/~brad/data/PKNL_BUGS.txt
- Interstim multivariate model:  
  www.biostat.umn.edu/~brad/data/InterStim.odc
- Bayesian $p$-values (illustrated with stacks data):  
  www.biostat.umn.edu/~brad/data/stackspval_BUGS.txt

**BRugs**
- Design for binary and Cox PH models: Brian Hobbs’ webpage:  
  www.biostat.umn.edu/~brianho/papers/2007/JBS/prac_bayes_design.html

*Thanks for your attention!*