Different approximations and methods for calculating the FIM and their consequences

Joakim Nyberg, Joakim Ringblom, Mats O. Karlsson, Andrew C. Hooker

Division of Pharmacokinetics and Drug Therapy
Department of Pharmaceutical Biosciences
Uppsala University
Sweden

2008-06-23
Background

Several software are available in the Population Optimal Design area and they all have different options for how to calculate the FIM and also how to include residual variability in the model.

How can this effect the optimal design?
Introduction

• The Population Fisher Information Matrix (FIM) can be very time consuming to optimize over => The reduced FIM is quite often used, what consequences does that have on the optimization?

• The residual variance of a model is often proportional + additive (slope+const), how will different implementations of the variance model affect the optimal design?

• The FIM is often parameterized differently in terms of SD or variances, consequences?!
Full vs. Reduced FIM - Example

\[
FIM^{-1}_{\text{Full}} = \begin{pmatrix}
\theta_{CL} & \theta_V & \omega_{CL}^2 & \omega_V^2 \\
\theta_{CL-V} & \theta_{CL-V} & \theta_{CL-\omega_{CL}^2} & \theta_{CL-\omega_V^2} \\
\theta_{CL-\omega_{CL}^2} & \theta_{CL-\omega_{CL}^2} & \theta_{CL-V} & \theta_{CL-V} \\
\theta_{CL-\omega_V^2} & \theta_{CL-\omega_V^2} & \theta_{CL-V} & \theta_{CL-V} \\
\end{pmatrix}
\]

\[
FIM^{-1}_{\text{Reduced}} = \begin{pmatrix}
\theta_{CL} & \theta_{CL-V} & \omega_{CL}^2 & \omega_{CL-V}^2 \\
\theta_{CL-V} & \theta_{CL-V} & \omega_{CL-V}^2 & \omega_{CL-V}^2 \\
\omega_{CL}^2 & \omega_{CL-V}^2 & \omega_{CL-V}^2 & \omega_{CL-V}^2 \\
\omega_{CL-V}^2 & \omega_{CL-V}^2 & \omega_{CL-V}^2 & \omega_{CL-V}^2 \\
\end{pmatrix}
\]
Different models investigated with Full vs. Reduced FIM

- First order absorption model
- Bateman function
- Biorhythm model
- Derendorf surge model
- Enterohepatic recirculation model
- IV PK/Emax PD model
- Lag-time model
- Michalis-Menten elimination model
- Pool-tolerance model
- Transit compartment absorption model
- Viral load model
HIV Viral load

$$E_{\text{group 1}} = \log_{10}\left( e^{p_1} \cdot e^{-e^{d_{1,t}}} + e^{p_2} \cdot e^{-e^{d_{2,t}}} \right)$$

$$E_{\text{group 2}} = \log_{10}\left( e^{p_1} \cdot e^{-e^{(d_{1,t}+\beta)t}} + e^{p_2} \cdot e^{-e^{d_{2,t}}} \right)$$

HIV Viral load model

Reduced FIM

= optimal value
HIV Viral load model

Full FIM

= optimal value
Derendorf surge model

\[ y = R_0 \cdot \left( 1 + \frac{SA}{\left( \frac{t - T_0}{SW} \right)^N + 1} \right) \]
Derendorf surge model-OD

Optimal sample times for Derendorf surge model with full IM

Optimal sample times for Derendorf surge model with reduced IM

Sampling time 2 in group 1

Sampling time 3 in group 1

Sampling time 2 in group 1

Sampling time 3 in group 1
Biorhythm & Transit compartment

\[
A = \frac{k_{in}}{k_{out}} + \frac{\text{amp} \cdot k_{in}}{k_{out}^2 + \left(\frac{2\pi}{\text{per}}\right)^2} \left[ k_{out} \cdot \cos\left(\left(t - \text{peak}\right) \cdot \frac{2\pi}{\text{per}}\right) + \frac{2\pi}{\text{per}} \cdot \sin\left(\left(t - \text{peak}\right) \cdot \frac{2\pi}{\text{per}}\right) \right]
\]

\[
\frac{\partial A_{abs}}{\partial t} = e^{\frac{A_n}{n+1} \cdot \text{MTT}} - k_a \cdot A_{abs}
\]

\[
\frac{\partial A_{central}}{\partial t} = k_a \cdot A_{abs} - \frac{CL}{V} \cdot A_{central}
\]
Biorhythm results

- Singular FIM -> Hard/Impossible to optimize on!
- Similar for Transit compartment model
Why singularity for reduced?

Biorhythm:

Rows of \( M_1 = \frac{\partial f(\bar{x}_i, \Theta)}{\partial \theta} \) are linearly dependent

Transit model:

Rows of \( M_3 = \frac{\partial \text{vec}(\text{Var}(\bar{y}_i))}{\partial \omega^2}, \frac{\partial \text{vec}(\text{Var}(\bar{y}_i))}{\partial \sigma^2} \) are linearly dependent
Similar optimal designs with reduced FIM

- First order absorption model
- Bateman function model
- Enterohepatic recirculation model
- IV PK/Emax PD model
- Lag-time model
- Michaelis-Menten elimination model
- Pool-tolerance model
Residual Variability models

General model

\[ \bar{y}_i = f\left(\bar{\theta}_i, \bar{x}_i\right) + h\left(\bar{\theta}_i, \bar{x}_i, \bar{\epsilon}_i\right) \]
\[ \bar{\epsilon}_i \sim N(0, \Sigma) \]
\[ \bar{\eta}_i \sim N(0, \Omega) \]

Additive + proportional residual variability model

\[ \sigma_{add}^2 + \sigma_{prop}^2 \cdot \left(f\left(\bar{\theta}, \bar{x}\right)\right)^2 \]
\[ \bar{\epsilon}_i \sim N(0, \Sigma) \]

Additive + proportional residual variability model

\[ \left(\sigma_{add} + \sigma_{prop} \cdot f\left(\bar{\theta}, \bar{x}\right)\right)^2 \]
\[ \bar{\epsilon}_i \sim N\left(0,\Sigma\right) \]
General Variance

General model

$$\tilde{y}_i = f(\tilde{\theta}_i, \tilde{x}_i) + h(\tilde{\theta}_i, \tilde{x}_i, \tilde{\epsilon}_i)$$

$$\tilde{\epsilon}_i \sim N(0, \Sigma)$$

$$\tilde{\eta}_i \sim N(0, \Omega)$$

General variance

$$\text{var}(\tilde{y}_i) = L \cdot \Omega \cdot L^T + \text{diag}(H \cdot \Sigma \cdot H^T)$$

$$L_i(\tilde{x}_i, \tilde{\theta}_i) \equiv \left. \frac{\partial f}{\partial \tilde{\eta}_i}(\tilde{x}_i, \tilde{\theta}_i) \right|_{\tilde{\eta}_i=0}$$

$$H_i(\tilde{x}_i, \tilde{\theta}_i) \equiv \left. \frac{\partial h}{\partial \tilde{\epsilon}_i}(\tilde{x}_i, \tilde{\theta}_i, \tilde{\epsilon}_i) \right|_{\tilde{\eta}_i=0, \tilde{\epsilon}_i=0}$$
General Variance applied to different residual models

$$\sigma_{\text{add}}^2 + \sigma_{\text{prop}}^2 \cdot f(\bar{\theta}, \bar{x})^2$$

$$\text{var}(\tilde{y}_i) = \text{diag}(H \cdot \Sigma \cdot H^T) = \text{diag} \left( \begin{pmatrix} 1 & f(t_1, \bar{\theta}) \\ 1 & f(t_2, \bar{\theta}) \end{pmatrix} \cdot \begin{pmatrix} \sigma_{\text{add}}^2 & 0 \\ 0 & \sigma_{\text{prop}}^2 \end{pmatrix} \cdot \begin{pmatrix} 1 & f(t_1, \bar{\theta}) \\ 1 & f(t_2, \bar{\theta}) \end{pmatrix}^T \right) =$$

$$\begin{pmatrix} \sigma_{\text{add}}^2 + \sigma_{\text{prop}}^2 \cdot f(t_1)^2 & 0 \\ 0 & \sigma_{\text{add}}^2 + \sigma_{\text{prop}}^2 \cdot f(t_2)^2 \end{pmatrix}$$

$$\left( \sigma_{\text{add}} + \sigma_{\text{prop}} \cdot f(\bar{\theta}, \bar{x}) \right)^2$$

$$\text{var}(\tilde{y}_i) = \text{diag}(H \cdot \Sigma \cdot H^T) = \text{diag} \left( \begin{pmatrix} 1 & f(t_1, \bar{\theta}) \\ 1 & f(t_2, \bar{\theta}) \end{pmatrix} \cdot \begin{pmatrix} \sigma_{\text{add}}^2 \sigma_{\text{add}} \cdot \sigma_{\text{prop}} \sigma_{\text{prop}}^2 \end{pmatrix} \cdot \begin{pmatrix} 1 & f(t_1, \bar{\theta}) \\ 1 & f(t_2, \bar{\theta}) \end{pmatrix}^T \right) =$$

$$\begin{pmatrix} \left( \sigma_{\text{add}} + \sigma_{\text{prop}} \cdot f(t_1) \right)^2 & 0 \\ 0 & \left( \sigma_{\text{add}} + \sigma_{\text{prop}} \cdot f(t_2) \right)^2 \end{pmatrix}$$
Implications of residual models

\[ \sigma^2_{\text{add}} + f^2 \sigma^2_{\text{prop}} \]

\( \text{CV: } \theta_1 = 0.9 \quad \sigma^2_{\text{add}} = 1.6 \quad \theta_2 = 2.2 \quad \sigma^2_{\text{prop}} = 5.2 \)

\[ (\sigma_{\text{add}} + f \sigma_{\text{prop}})^2 \]

\( \text{CV: } \theta_1 = 1.3 \quad \sigma^2_{\text{add}} = 2.5 \quad \theta_2 = 2.8 \quad \sigma^2_{\text{prop}} = 5.9 \)
Different derivations of FIM, w.r.t. variance vs. stdev

Variance derivation
\[ \frac{\partial \text{var}(\bar{y})}{\partial \sigma^2} = H \cdot H^T \quad \rightarrow \quad SE_{\sigma^2} = 2 \cdot SE_\sigma \]

Stdev derivation
\[ \frac{\partial \text{var}(\bar{y})}{\partial \sigma} = 2 \cdot \sigma \cdot H \cdot H^T \quad \rightarrow \quad SE_\sigma = \frac{SE_{\sigma^2}}{2} \]

The difference in SE will reflect differences in |FIM| and vice versa
Conclusions

• It is important to increase the understanding of potential differences between software in OD.

• Different approximations and different residual models can clearly affect the optimal design and in some cases lead to different results.

• When comparing the expected uncertainty of an estimator, the residual variability model needs to be considered to get an accurate comparison.

• Allowing for a general error function and linearizing around it is the most flexible way to allow for all combinations of models and correlations.
Backup - Calculating FIM

\[ \text{FIM}(\Theta)_i = \begin{pmatrix} M_{1i} & 0 \\ M_{2i} & M_{3i} \end{pmatrix}^T \left( \text{Var}(\tilde{y}_i)^{-1} \begin{pmatrix} 0 \\ M_{4i}^{-1} \end{pmatrix} \begin{pmatrix} M_{1i} & 0 \\ M_{2i} & M_{3i} \end{pmatrix} \right) \]

\[ M_{1i} = \frac{\partial f(\tilde{x}_i, \tilde{\theta})}{\partial \tilde{\theta}} \]

\[ M_{3i} = \begin{bmatrix} \frac{\partial \text{vec}(\text{Var}(\tilde{y}_i))}{\partial \tilde{\omega}^2} & \frac{\partial \text{vec}(\text{Var}(\tilde{y}_i))}{\partial \tilde{\sigma}^2} \end{bmatrix} \]

\[ M_{2i} = \frac{\partial \text{vec}(\text{Var}(\tilde{y}_i))}{\partial \tilde{\theta}} \]

\[ M_{4i} = 2 \cdot \text{Var}(\tilde{y}_i) \otimes \text{Var}(\tilde{y}_i) \]

\[ \text{Var}(\tilde{y}_i) \approx L \cdot \Omega \cdot L^T + \text{diag}(H \cdot \Sigma \cdot H^T) \]