Shrinkage in Empirical Bayes Estimates for Diagnostics and Estimation: Problems and Solutions

Radojka M. Savić & Mats O. Karlsson

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

✓ Empirical Bayes Estimates
✓ Use in Non-linear Mixed Effects Modelling
✓ Shrinkage phenomenon
✓ Shrinkage related problems:
  Diagnostics
  Estimation process (FOCE & NONP)
✓ Solutions & Recommendations
Empirical Bayes Estimates

POSTHOC estimates – individual parameter estimates

Provide population PKPD modellers with:

✓ EBE - individual parameter estimate
✓ IPRED – individual predictions
✓ IWRES – individual weighted residuals

\[ IWRES_{ij} = \frac{(DV_{ij} - IPRED_{ij})}{SD(\varepsilon_{ij})} \]
Use of EBEs

✓ **Diagnostics**
  - IPRED vs DV
  - IWRES vs IPRED
  - EBE vs EBE
  - EBE vs Covariate
  - GAM

✓ **Estimation**
  - FOCE
  - Nonparametric estimation

✓ **Prediction (TDM)**

✓ **Simulation**
Diagnostics based on EBEs

Increases resolution by separating variability components

If data are uniformative:

1. EBE distribution will shrink towards 0 (population mean)
   \[ \text{EBE} \rightarrow 0 \]

2. Individual predictions (IPRED) will shrink towards the corresponding observation (DV)
   \[ \text{IPRED} \rightarrow \text{DV} \]

3. IWRES, residual components will shrink towards 0
   \[ \text{IWRES} \rightarrow 0 \]

R.M. Savic, J.J. Wilkins and M.O. Karlsson.
Concept of EBE shrinkage

How well can we estimate this parameter?

η distribution

η

0

No data

Sparse data

Rich data

TRUE parameter

Rich data

η_\text{i}
Shrinking EBE distribution towards 0

Probability Density Function

Post Hoc $\eta$ values

- True distribution
- Rich data
- Sparse data
- Very sparse data

$\Omega$ decrease
Shrinking IPRED towards DV
Shrinking IPRED towards DV
Shrinking IPRED towards DV

IPRED is shrinking towards DV
Shrinking IWRES towards 0

\[ IWRES_{ij} = \frac{DV_{ij} - \text{IPRED}_{ij}}{SD(\varepsilon)} \]

If IPRED \xrightarrow{} DV

IWRES \xrightarrow{} 0

![Graph showing probability density function of IWRES values with true IWRES and shrunk IWRES distributions.](image)
Quantifying Shrinkage

1. ETA shrinkage

\[ \eta_{sh} = 1 - \frac{SD(\hat{\eta}_{ph})}{\omega} \]

2. EPSILON shrinkage

\[ \varepsilon_{sh} = 1 - SD(IWRES) \]

How do these values change with information content?
How shrinkage may influence diagnostics?

**Diagnostics explored:**

1. **EBE-related diagnostics** ($\eta$ - shrinkage)
   - EBE distribution plots
   - EBE vs EBE plots
   - EBE vs Covariate plots

2. **IPRED / IWRES - related diagnostics** ($\varepsilon$ - shrinkage)
   - IPRED vs DV plot
   - IWRES vs IPRED plot

**Methods:** MC simulations
- True model was fitted to data unless otherwise stated
- Graphical diagnostics showed on single simulation example to facilitate visualization
Consequences of \( \eta \)- shrinkage:

(i) Change of distribution shape

ETABAR is the arithmetic mean of the ETA-estimates, and the p-value is given for the null hypothesis that the true mean is 0.

ETABAR: \(-0.75E-01\) \(0.63E+00\) \(-0.30E+00\)
SE: \(0.34E-01\) \(0.17E+00\) \(0.17E+00\)

P VAL.: \(0.28E-01\) \(0.13E-03\) \(0.71E-01\)

Significant change in mean value of ETAs!
Consequences of \( \eta \)-shrinkage

(iii) parameter correlation (hidden)

- True data correlated (CORR=0.65)
- Estimated data Correlation hidden (CORR=0.2)

\( \eta_{sh} \) = 22%

\( \eta_{sh} \) = 48%
Consequences of $\eta$- shrinkage

(iii) parameter correlation (hidden)

<table>
<thead>
<tr>
<th>Correct</th>
<th>shrinkage 20-30%</th>
<th>shrinkage 30-40%</th>
<th>shrinkage 35-45%</th>
</tr>
</thead>
<tbody>
<tr>
<td>CL</td>
<td>CL</td>
<td>CL</td>
<td>CL</td>
</tr>
<tr>
<td><strong>Corr: 0.65</strong></td>
<td><strong>Corr: 0.38</strong></td>
<td><strong>Corr: 0.14</strong></td>
<td><strong>Corr: 0.02</strong></td>
</tr>
</tbody>
</table>

Correct shrinkage 20-30%
Consequences of $\eta$- shrinkage

(iii) parameter correlation (induced)

Commonly induced correlations:

- $\eta_{sh} = 28\%$
- $\eta_{sh} = 53\%$
- $ka \sim V$
- $EC_{50} \sim E_{max}$
- $EC_{50} \sim k_{out}$
Consequences of $\eta$- shrinkage

(iv) parameters / $\eta_s$ vs. Covariates

Covariate – weight (kg)
Consequences of $\eta$- shrinkage - Summary

(i) EBEs may, in addition to shrinkage, show non-normal distribution even when the underlying $\eta$ distribution is normal

(ii) mean values of EBEs ("ETABAR") may be significantly different from zero, even for a correctly-specified model (a result of asymmetric $\eta$-shrinkage)

(iii) EBE-EBE correlations may be hidden or even induced

(iv) covariate relationships may be hidden, falsely induced, or the shape of the true relationship distorted
Consequences of $\varepsilon$-shrinkage

(i) Low power of IPRED to detect model misspecification

Plot is a clear indication of model misspecification

Fitted model: first order absorption

True model: zero-order absorption model
Consequences of $\varepsilon$- shrinkage

(i) Low power of IPRED to detect model misspecification

"Perfect fit" phenomenon

$\varepsilon_{sh} = 6\%$

$\varepsilon_{sh} = 69\%$
Emax model fitted to data simulated with a sigmoidal Emax model

\[ \varepsilon_{sh} = 5\% \quad \varepsilon_{sh} = 13\% \quad \varepsilon_{sh} = 29\% \]

Consequences of $\varepsilon$- shrinkage

(ii) Low power of IWRES to diagnose residual error misspecification

11 obs/ID (3 etas)

4 obs/ID (3 etas)

Misspecification indicated

Misspecification NOT indicated
(ii) low power of IWRES to diagnose residual error model misspecification
Consequences of $\varepsilon$- shrinkage - Summary

(i) low power of IPRED to diagnose structural model misspecification ("perfect fit" phenomenon)

(ii) low power of IWRES to diagnose residual error model misspecification
Conclusions – part 1

✓ Model diagnostics involving EBE, IPRED, IWRES is misleading in the presence of shrinkage.

✓ The problem of shrinkage in showed examples associated to the diagnostics solely. Estimation is not affected.

✓ Consequences of shrinkage ignorance:
  - wrong decisions
  - increased time for data analysis
  - wrong models

✓ Shrinkage phenomenon is likely to affect other type of model diagnostics such as:
  - GAM
  - CWRES
1. **Report the shrinkage extent!**
   - Inform modelers about relevance of the graphs

2. **Estimate standard errors of ETAs**
   - Refine EBEs and EBE-based diagnostics

3. **Use other type of diagnostics**
   - Simulation based diagnostics

4. **Do more model testing inside NONMEM**
Background:
EBEs are computed at each iteration step

Question:
How shrinkage may affect FOCE method?
EBE shrinkage & FOCE

Bias in FOCE parameter estimates with EBE shrinkage

\[
\eta_{sh} < 10\%
10 < \eta_{sh} < 25\%
\eta_{sh} > 25\%
\]
Conclusions - part 2

1. Increased bias & variance of FOCE parameter estimates in the presence of shrinkage

→ **FOCE method is becoming more like FO method**
  - Biased variance estimates

**Solution**

✓ new algorithms
✓ refine variance estimates (NONP)
EBE shrinkage & $\text{NONP}$

1. Search for support points
   - parametric step (FO/FOCE)
   - EBEs computation
   - Points of support

2. Probability estimation
   - the joint probability
   - the marginal cumulative probability

What if EBEs are shrunk?
Consequences of EBE shrinkage

QQ plot: true versus estimated distribution

DEFAULT NONP distribution
NONP and EBE shrinkage
How to proceed?

1. **Keep using default NONMEM support points**
   - range of available support points may be sufficient
   - range of available support points lower than expected
     - results still may be improved compared to the parametric outcome

2. **Inflate variances prior to EBE (POSTHOC) estimation**
   (enough to inflate twice the variances)
Improvement with the inflated variances method
There is a way to enhance the NONP grid with additional points of support.

Additional points of support are generated via simulations from final model.

Practically it requires:

- simulation from the final model
- computation of the individual contributions to the entire NONP density

A general routine that automizes this is under development.
Improvement with the enriched grid method

**Diagram:**

- **TRUE distribution**
- **ESTIMATED distribution**

- **Default NONP distribution**
- **NONP based on enriched grid**
Conclusions

1. Model diagnostics involving EBE, IPRED, IWRES is misleading
   Essential part of model building:
   - wrong decisions
   - increased time for data analysis
   - wrong models

2. FOCE method is becoming more like FO method
   Biased variance estimates

3. NONP method may be biased
   At higher shrinkage extents
Take-home message

Compute the shrinkage!
1. **Model diagnostics**
   - Report the shrinkage extent!
   - Compute standard errors of ETAs
   - Use other type of diagnostics
   - More testing directly in NONMEM

2. **FOCE**
   - new algorithms
   - refine variance estimates (NONP)

3. **NONP method**
   - Inflate variances prior to EBE estimation
   - Use extended grid method (soon available in PsN)
Falsely induced covariate relationships by EBE shrinkage

Simulations:
- WT as covariate on V
- no covariance V-Ka
- no influence of WT on ka.