Full Covariate Models as an Alternative to Methods Relying on Statistical Significance for Inferences about Covariate Effects: A Review of Methodology and 42 Case Studies

PAGE 2011
Athens, Greece

Marc R. Gastonguay, Ph.D.
### Table III (continued)

<table>
<thead>
<tr>
<th>Model</th>
<th>Question</th>
<th>Equation</th>
<th>Objective Function</th>
<th>Conclusion (Significance = Yes)</th>
</tr>
</thead>
</table>
| Model 15 | Does condition (Healthy/asthmatic) affect the volume of distribution (V2) (LINEAR MODEL)? | \[
\text{TVKA} = \theta (1) \\
\text{TVCL} = \theta (2) \\
\text{TVV2} = \theta (3)*\text{COND} + \theta (4)*(1 - \text{COND}) \\
\theta (1) = 4.42 \text{ h}^{-1} \\
\theta (2) = 392 \text{ L/h} \\
\theta (3) = 1170 \text{ L} \\
\theta (4) = 1510 \text{ L} \\
\] | $-9603$ | No |
| Model 16 | Does condition (Healthy/asthmatic) affect the clearance (POWER MODEL)? | \[
\text{TVKA} = \theta (1) \\
\text{TVCL} = \theta (2)*\theta (3)**\text{COND} \\
\text{TVV2} = \theta (4) \\
\theta (1) = 4.42 \text{ h}^{-1} \\
\theta (2) = 363 \text{ L/h} \\
\theta (3) = 468 \\
\theta (4) = 1720 \text{ L} \\
\] | $-240$ | No |
| Model 17 | Does condition (Healthy/asthmatic) affect the clearance (LINEAR MODEL)? | \[
\text{TVKA} = \theta (1) \\
\text{TVCL} = \theta (2)*\text{COND} + \theta (3)*(1 - \text{COND}) \\
\text{TVV2} = \theta (4) \\
\theta (1) = 4.42 \text{ h}^{-1} \\
\theta (2) = 384 \text{ L/h} \\
\theta (3) = 438 \text{ L/h} \\
\theta (4) = 1180 \text{ L} \\
\] | $-9598$ | No |
| Model 18 | Does weight affect clearance and volume of distribution (POWER MODEL)? | \[
\text{TVCL} = \theta (2)*\text{(WT/75)}**\theta (3) \\
\text{TVV2} = \theta (4)*\text{(WT/75)}**\theta (5) \\
\theta (2) = 393 \text{ L/h} \\
\theta (3) = 0.56 \\
\theta (4) = 1220 \text{ L} \\
\theta (5) = 1.28 \\
\] | $-9738$ | Yes |
Appropriate Inference?

Covariate Effects:
Creatinine clearance was the most significant covariate on clearance. The clinical importance of this finding is unknown, since less than 2% of dose is excreted in the urine. This may be an artifact of the data as the current analysis data set did not include patients with moderate or severe renal impairment. Weight, age and sex were not significant covariates and, therefore, require no dose adjustment.
Full Models


- Harrell, F.E. Regression Modeling Strategies. 2001; Springer-Verlag. NY.

Overview

- Objectives of Covariate Model Building
- Problems with Traditional Stepwise Methods
- Full Covariate Models
  - Data Reduction
  - Model Development
  - Inferences about Covariate Effects
- Review of Case Studies
- Aligning Covariate Modeling Methods with Objectives
- Summary
Objectives of Covariate Model Development

- Explain “random” variability in parameters and response

- Understand causes of variability and apply the knowledge
  - For better clinical therapeutic use (dosing, adjustment, labeling)
  - To allow for better control in clinical trials
  - In other words, make inferences about covariate effects from modeling results

- Improve predictive performance of the model
  - For subjects in the current data set
  - For trial simulation of future studies
  - For future patient populations
\[ TVCL = \theta_1 + \theta_2 \times CRCL \]

\[ CL_i = TVCL + \eta_i \]
Traditional Covariate Screening Methods in Pop PKPD

Outside of population model context:
- Exploratory Graphics of Individual Random Effects ($\eta$) vs. Covariates
- Generalized Additive Modeling

Within population model context:
- Stepwise Forward Addition
- Stepwise Backward Elimination
- Stepwise Forward/Backward

These are problematic in presence of: $\eta$-shrinkage, imbalanced designs, time-dependent covariates, plus other problems…

Other problems with these methods… more later.
Stepwise Backward Elimination

What happens when a covariate effect is statistically "significant", but not clinically important?

If a covariate effect is not statistically "significant", does this mean that there is no effect of that covariate?

**Step 1**
Start with full model

- OFV (Full)
  \[ CL = \theta_1 + \theta_3 \text{AGE} + \theta_5 \text{WT} \]
  \[ V = \theta_2 + \theta_4 \text{WT} \]

**Step 2**
Set each \( \theta \) to null value and record OFV for each run

If any are not significant \( p<0.001 (\Delta \text{OFV} <10.88) \) remove covariate with smallest \( \Delta \text{OFV} \)

Repeat stepwise until all covariates are significant at \( p<0.001 \)

- "REDUCED MODEL"
- Refine variance model structure and parameters.
  - Run \$COV

- "FINAL MODEL"
Some Problems with Stepwise Regression

- Based on methods (e.g. F tests for nested models) that were intended to be used to test pre-specified hypotheses

- p-values are difficult to interpret and difficult to adjust appropriately for multiple comparisons

- Regression coefficients are typically over-estimated (e.g. selection bias, false-positive findings, or biased inference).

- Confidence intervals are falsely narrow.

- Severe problems in the presence of correlated or collinear predictors (estimation bias, interpretation difficulties).

- Resulting models may be predictive of the current data set, but often are difficult to interpret or generalize.
More Problems with Stepwise Regression

- NONMEM likelihood approximations can result in incorrect p-values, even when model is known.

- Reconciling statistically significant effects with clinically important effects is challenging.

- Lack of statistical significance does not necessarily indicate lack of effect.

- Even with very large data sets, and rigorous model building, testing and cross-validation (conditions not typically seen in population PKPD), stepwise selection often fails.
A Purpose-Driven Parsimony Principle

“When competing hypotheses are equal in other respects, select the hypothesis that introduces the fewest assumptions and postulates the fewest entities while still sufficiently answering the question.” – Occam's razor

- Stepwise p-value reduced models do not allow for inferences about “non-significant” covariate effects and result in biased standard errors and point estimates. They do not sufficiently answer the question about clinical importance of covariate effects.

- For the purpose of making inferences about covariate effects, the full covariate model is the most parsimonious model.
Parsimony Principle Restated

“All things *being equal*, choose the simpler model.”
- Unknown

"We are to admit no more causes of natural things than such as are both *true and sufficient* to explain their appearances.”
- Isaac Newton

“Make everything as simple as possible, *but not simpler.*”
- Albert Einstein

“The simplest explanation that *covers all the facts* is usually the best.”
- Unknown
Full Covariate Model Approach

- Define stable base model structure based on GOF criteria.

- Data Reduction:
  - Avoid searching across all possible covariates
  - Avoid correlated predictors
  - Rational selection of potential covariates for full model

- Re-parameterize as necessary to develop a stable full model.

- Estimate all parameters of full model and construct intervals or posterior distributions (bootstrap, Bayesian, cov-matrix of est.).

- Make inferences based on posterior predictive intervals of estimated covariate effects

- Explore remaining trends as secondary hypothesis-generating step
Before Covariate Model Building: Data Reduction

- Examine covariate data
  - Identify range & distribution of continuous covariates
  - Count number in each category for categorical covariates
  - Identify strong correlations or collinearity between covariates
    - Select covariates with unique information
    - May require composite/interaction to convert to single variable

- Was the study designed to estimate covariate effects?

- How do inclusion criteria impact choice of covariates to include in modeling?

- Does covariate make sense given prior knowledge?

- Which covariates are of interest from clinical perspective?
Correlation/ Collinearity

Covariate effects to be included in model should be independent, e.g. they carry unique information.

Rule of thumb:
Be cautious when |corr. coef.| > 0.3
Check for Independence of Covariates

- Relationships between continuous/categorical covariates
  - Explore graphically
  - Apply ANOVA and/or Kruskal-Wallis test
Solutions to Correlation / Collinearity

- Avoid simultaneous inclusion of suspect covariates
- Remove correlation
  - For example: MDRD calculation for renal function is normalized by BSA, and can be included simultaneously with measures of body size
- Seek additional data where the same variables are independent
  - Include data from renal impairment study where CRCL~WT are not likely to be correlated
- Create a single summary variable to represent correlated predictors
  - BMI reflects both weight and height
- Fix one of the covariate-parameter relationships
  - Age and weight are highly correlated in pediatrics, but fixing weight relationship to an allometric expression allows estimation of age effects
- Reserve correlated covariates for exploratory modeling
Guiding Factors for Selection of Covariates for FCM

- Perform data reduction step

- Create focused questions about specific covariate effects in the current data set, based on:
  - Scientific or clinical interest
  - Mechanistic plausibility
  - Prior knowledge about covariate effects

(These should be defined \textit{a priori} in the analysis plan).

- Pre-specified covariate selection is not subject to problems of eta-shrinkage or data-driven selection bias
## Pre-Specified Covariate Plan

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Model Parameters</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight</td>
<td>CL, V1, Q, V2</td>
<td>Clinical interest</td>
</tr>
<tr>
<td>Age</td>
<td>CL</td>
<td>Clinical interest</td>
</tr>
<tr>
<td>Race</td>
<td>CL, V1</td>
<td>Clinical interest. Bridging goal.</td>
</tr>
<tr>
<td>Disease State Type</td>
<td>CL</td>
<td>Clinical interest</td>
</tr>
<tr>
<td>Child-Pugh Score</td>
<td>CL</td>
<td>Clinical interest. Prior knowledge of hepatic elimination mechanism; CYP3A4</td>
</tr>
<tr>
<td>Drug X Interaction</td>
<td>CL</td>
<td>Clinical interest. Known CYP3A4 inhibitor and common con-med</td>
</tr>
<tr>
<td>etc...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Stable Parameterization of Full Model

an example:

\[ TVP = \theta_n \cdot \prod_{1}^{m} \left( \frac{cov_{mi}}{ref_m} \right)^{\theta(m+n)} \cdot \prod_{1}^{p} \theta^{cov_{pi}}_{(p+m+n)} \]

where: the typical value of a model parameter (TVP) is described as a function of \( m \) individual continuous covariates (\( cov_{mi} \)) and \( p \) individual categorical (0-1) covariates (\( cov_{pi} \)) such that \( \theta_n \) is an estimated parameter describing the TVP for an individual with covariates equal to the reference covariate values (\( cov_{mi} = ref_m, cov_{pi} = 0 \)); \( \theta_{(m+n)} \) and \( \theta_{(p+m+n)} \) are estimated parameters describing the magnitude of the covariate-parameter relationships.
First, define effect magnitude likely to be clinically important (e.g. greater than +/- 20% of null value)

**Clinically Important:** Entire 95% interval of posterior distribution for covariate effect lies within clinically important region (always SS)

**Not Clinically Important:** Entire 95% interval of posterior distribution for covariate effect lies within clinically unimportant region. May be important in combination with other effects. (NSS or SS)

**Insufficient Information:** 95% interval of posterior distribution for covariate effect spans across values of covariate effect that are both clinically important and unimportant. (NSS or SS)

**Or... Probabilistic Approach:** Quantitatively describe probability of being clinically important using posterior distribution and reference range.
- Requires some measure of covariate effect parameter precision.

- Magnitude of covariate effect and precision of the estimate are key drivers of inferences, relative to clinically important effect size.

- Example: multiplicative binary categorical covariate effect (TVP=$\theta_1 \times \theta_2^{\text{cat}}$)
Check Full Covariate Model for Remaining Trends

- Plot CWRES or $\eta$ from Full Model vs. any covariates in database.
- and/or -
- Proceed with exploratory stepwise regression, starting at the full model. (For hypothesis-generating purposes).
Reduction of Full Model for Predictive Purposes

- Drop covariate effects that meet *both* of these criteria:
  - Not statistically significant (e.g. C.I. includes null value)
  - Not clinically important (e.g. entire C.I. is contained within no effect range)

- Retain all other effects (any one or more of these criteria):
  - Clinically important
  - Statistically significant (e.g. C.I. excludes null value)
  - Not statistically significant, but may be clinically important (e.g. characterized by insufficient information with C.I. extending into ranges of potential clinical importance)

*Removal of covariate effects should not impact coefficients for other remaining effects*
Some Examples
Population PK Model

Covariate effects on varenicline PK parameters

Typical individual: white, male, 45 years, CLcr = 100 mL/min, WT = 70 kg

- Typical CL/F = 10.4 L/hr
- CLcr (16-150 mL/min)
  - CLcr 20 mL/min
  - CLcr 40 mL/min
  - CLcr 65 mL/min
- Race = Black
- Race = Other
- Typical V2/F = 337 L
- Weight (40–130 kg)
- Age (18–75 yrs)
  - Race = Black
  - Race = Other
- Typical V3/F = 78.1 L
- Weight (40–130 kg)
- Typical Q/F = 2.08 L/hr
- Weight (40–130 kg)
Population PD Model

Reference: white, male, 45 years, first cigarette within 6–30 min

(Pop = 0.562)

Female
First cigarette ≥60 min
First cigarette within 31–60 min
First cigarette within <5 min
Age = 18 years
Age = 75 years
Black
Other races

Ratio of probability (W9-12 CAR) in subpopulation to the probability (W9-12 CAR) in reference population
Probability-Based Assessment

Quantify probability of covariate effect being clinically important.

Numbers indicate percent of posterior probability distribution relative to reference region.
Expected Covariate Impact Across Multiple Effects

Posterior simulation (e.g. simulation with parameter uncertainty) of quantity of interest: posterior probability distribution of AUC\textsubscript{24ss} vs. weight, conditioned by age (55 or 80 years), patient type, and degree of renal impairment.
Review of FCM Case Studies

- 42 population analyses (32 PK, 10 PD) using NONMEM® v. 5, 6.x or 7.x
- All late stage development (end of Phase 2b, or Phase 3)
- 95% CI of parameter estimates obtained from stratified non-parametric bootstrap (95% of cases) or NONMEM® asymptotic standard errors (5% of cases)
- Statistical significance was defined as exclusion of null value in 95% CI
- Clinical importance was defined based on 95% CI relative to clinical no-effect range (e.g. null value +/- 20%) for AUC (PK) or a PD response parameter.
Results of FCM Case Studies

- Models converged successfully: 100%
- Successful $COV$ step: 98%
- Number of covariates in source data: 14 (median), 4-60 (range)
- Number of covariates in FCM: 6 (median), 1-16 (range)

- Total number of covariate effects estimated: 258
  - Statistically significant covariates: 48%
  - Not statistically significant covariates: 52%
  - Clinically important covariate effects: 24%
  - Not clinically important effect: 48%
  - Insufficiently informed effect estimate: 28%
Results of Covariate Effect Inferences for Case Studies

- Insufficient Information
- Not Clinically Important
- Clinically Important
- Covariate must be in model to make inference

- Inferences are sensitive to biases typically associated with stepwise methods:
  > effect magnitude (falsely increased)
  > precision (falsely narrow confidence intervals)

- Other issues mentioned earlier…
Align Methods with Modeling Purpose

- **Hypothesis Generation / Data Mining**
  - Stepwise methods are useful for exploratory goals

- **Inferences Based on Estimation of Covariates Effects**
  - Use full covariate model

- **Prediction**
  - Stepwise models are useful, with large enough data set, and adequate test data sets, when selection bias is likely to be low; still be careful to include data reduction step
  - Full covariate models result in more appropriate (larger) parameter uncertainty, but can lead to larger prediction error of response. Empirical Bayes estimates of individual model parameters are unaffected.

- **Hypothesis Testing**
  - Requires *a priori* adequately powered design and model specification
  - Stepwise methods are not always appropriate
  - Limit model to specific hypothesis to be tested (e.g. pre-defined)
Summary

- Statistical significance does not predict clinical importance and should not be used for covariate effect inferences.

- Inferences about clinical importance are driven by estimated magnitude of effect and associated precision, and are sensitive to biases in these metrics.

- Stepwise p-value driven approaches and FCM approaches both have utility in modeling and simulation:
  - Earlier stage exploratory data mining (stepwise p-value)
  - Later stage clinical inferences (FCM)
“The data analyst knows more than the computer… failure to use that knowledge produces inadequate data analysis”.

Acknowledgements

- Tim Bergsma
- Nathanael Dirks
- Jeannine Fisher
- Leonid Gibiansky
- Bill Gillespie
- Bill Knebel
- John Mondick
- Matthew Riggs

- Ahmed El-Tahtawy
- Jonathan French
- Ken Kowalski
- Patanjali Ravva
- Industry/Academic Collaborators
- NMUSERS participants
Related Abstracts at PAGE 2011

I-11 A novel covariate search method intended for PKPD models with nonparametric parameter distributions. Paul G. Baverel (1), Radojka M. Savic (1), Scott F. Marshall (2), Mats O. Karlsson (1)

II-47 Selection Bias in Pre-Specified Covariate Models. Vijay D Ivaturi, Andrew C Hooker, Mats O Karlsson

II-53 Comparison of methods for handling missing covariate data. Åsa M. Johansson, Mats O. Karlsson

II-57 Evaluation of Stepwise Covariate Model Building Combined with Cross-Validation. Takayuki Katsube (1, 2), Akash Khandelwal (1), Kajsa Harling (1), Andrew C Hooker (1), Mats O Karlsson (1)

II-58 The bootstrap of Stepwise Covariate Modeling using linear approximations. Ron J Keizer, Akash Khandelwal, Andrew C Hooker, Mats O Karlsson


References (1)


Han, Phey Yen and Kirkpatrick, Carl M J and Green, Bruce. Informative study designs to identify true parameter-covariate relationships (article). J Pharmacokinet Pharmacodyn. 2009. 36 (2), 147-63.

Harrell, F.E. Regression Modeling Strategies. 2001; Springer-Verlag. NY.
References (2)


Wade JR, Beal SL, Sambol NC. Interaction between structural, statistical, and covariate models in population pharmacokinetic analysis. J Pharmacokinet

