Slide 1

Principles of Covariate Modelling

Nick Holford

Dept of Pharmacology & Clinical Pharmacology
University of Auckland
New Zealand

Slide 2

What is a Covariate ?

- "A covariate is any variable that is specific to an individual and may explain PKPD variability"
- Most important covariates are weight, renal function and age (in babies and infants)
- Examples of covariates that have been used in PKPD analysis

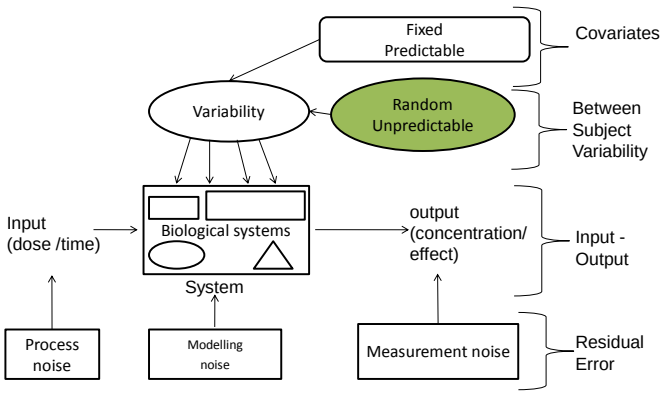
1. Size e.g. weight, fat free mass	11. Formulation
2. Renal disease e.g. Renal function	12. Diurnal variation
3. Age	13. Liver disease e.g. presence of hepatitis
4. Race	14. Genotype
5. Sex	17. Disease stage
6. Concomitant drug administration	18. Decontamination procedures
7. Clinical chemistry values e.g. bilirubin etc	
8. Hematologic values e.g. WBC count, hematocrit	
9. Protein Binding	

Covariates are used to describe predictable sources (fixed effects) of variability. A useful covariate is expected to explain some of overall variability and should lead to a decrease in unpredictable (random effects) variability. Covariate list suggested by Steve Duffull, University of Otago.

Slide 3

Covariate Models

- Types of covariates
- Types of covariate models
- Assessing influence of covariates
- Selecting the best covariate model

Slide 4	**Types of Covariates** • Classification 1 – Intrinsic factors (e.g. weight, age, sex, race, renal function, genotype) – Extrinsic factors (e.g. dose, compliance, smoker) • Classification 2 – Continuous (e.g. weight, age, renal function, dose, compliance) – Categorical (e.g. sex, race, genotype, smoker) 	Classification 1 is sometimes found in regulatory documents (e.g. FDA). Classification 2 is more relevant for using covariates with NONMEM.
Slide 5	**Why Model Covariates?** • To identify and explain between subject variability • Predict subject specific differences – By association with parameters • E.g. Volume of distribution increases linearly with weight `V=THETA(1)+THETA(2)*WT` $C = \frac{D}{V} \times \exp\left(-\frac{CL}{V} \times TIME\right)$ 	
Slide 6	**Where Do Covariates Fit in Population Analysis?**  	

Slide 7

How Should We Add Covariates?

1. Try every permutation in every order, set the runs going, come back later
   - … much later
   - … much much later
2. Choose your most important covariates first – but define your metric for importance…
3. What is important? Biology

Slide 8

Selecting Covariates

- Biological plausibility: Does the covariate have a biologically plausible explanation?
- Extrapolation plausibility: Does the model extrapolate sensibly outside the range of observed covariates?
- Clinical relevance: Is the covariate effect size clinically important?
- Statistical plausibility: Is the covariate statistically significant?

Slide 9

Biological Plausibility

- Continuous covariates
  - Linear model

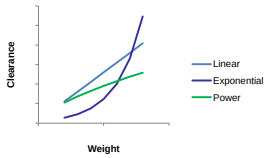
```
CL=THETA(1)+THETA(2)*WT/70
```

  - Exponential model

```
CL=THETA(1)*EXP(THETA(2)*WT/70)
```

  - Power model

```
CL=THETA(1)*(WT/70)**THETA(2)
```



- Which of these models is biologically plausible?
- Which models extrapolate sensibly?

More complex empirical models for allometry have been proposed by using WT in the exponent e.g. **CL=THETA(1)*WT**(THETA(2)*WT)**

Slide 10	## Standardizing Population Parameters $$CL = CL_{STD} \cdot \left(\frac{WT}{WT_{STD}} \right)^{3/4}$$ • The choice of WT_{STD} does not change the goodness of fit. So always use the widely used standard of 70 kg. • It becomes hard to compare parameters across studies if WT_{STD} is not consistent and especially if WT_{STD} is not defined! 	Holford NH. A size standard for pharmacokinetics. Clin Pharmacokinet 1996; 30: 329-32. Anderson BJ, Holford NHG. Tips and traps analyzing pediatric PK data. Pediatric Anesthesia 2011; 21: 222-37.
Slide 11	## Model Selection • Use the likelihood ratio test (LRT) – Difference between NONMEM Objective Function Value (OFV) • Do not use Empirical Bayes Estimates – See Karlsson & Savic 2007 for reasons why • Confirm covariates by backwards deletion from full model to obtain the final model • Use simulation based methods e.g. VPC to show importance of including covariate 	Karlsson MO, Savic RM. Diagnosing model diagnostics. Clin Pharmacol Ther 2007; 82: 17-20.
Slide 12	## Likelihood Ratio Test • The LRT compares two likelihoods • For nested models the difference of two log likelihoods is asymptotically chi-squared distributed: – For 1 parameter difference (1 degree of freedom) the critical value from the chi-squared distribution is '3.84' (P=0.05) – NONMEM has only approximate likelihoods so do not rely on P values as being 'exact' (Wählby et al. 2002) 	Wählby U, Bouw R, Jonsson EN, Karlsson MO. Assessment of Type I error rates for the statistical sub-model in NONMEM. Journal of Pharmacokinetics and Pharmacodynamics 2002; 29: 251-69.

Slide 13

Code for Covariate Models using NM-TRAN

Slide 14

The Population Value of CL [POPCL]

THETA(1) is the population value for clearance in the population

```
POPCL = THETA(1)
CL = POPCL*EXP(ETA(1))
```

However, the population value for CL can be refined, depending on some characteristic of the subjects

Slide 15

The Group Value of CL [GRPCL]

Here we define the population standard value for a 70 kg person and a group value for a specific WT

```
POPCL = THETA(1) ; L/h/70kg
GRPCL = POPCL*(WT/70)**0.75
```

Slide 16	**The Individual Value of CL [CL]**
	
	Finally we use a random effect model to describe the individual value
	
	`POPCL = THETA(1) ; L/h/70kg`
	`GRPCL = POPCL*(WT/70)**0.75`
	`CL = GRPCL*EXP(ETA(1))`
	
		Most biological parameters are non-negative. A normal distribution for a parameter such as CL is unsuitable because it includes negative values. The random effect ETA in NONMEM is always assumed to be normally distributed. A log-normal distribution of CL is expressed by exponentiating ETA.

Slide 17	**Allometric Scaling**
	
	In biology, functional (e.g. blood flow) and structural characteristics (e.g. tissue mass) scale predictably with allometric functions of WT. Thus CL, V can be modelled as:
	
	`POPCL = THETA(1)`
	`GRPCL = POPCL*(WT/70)**(3/4)`
	
	`POPV = THETA(2)`
	`GRPV = POPV*(WT/70)**1`
	`= POPV*WT/70`
	
	• Apply allometric weight models to all PK parameters
	• It is foolish to rely on statistical testing to decide if 'weight is in the model'
	
		Savage VM, Gillooly JF, Woodruff WH, West GB, Allen AP, Enquist BJ, Brown JH. The predominance of quarter-power scaling in biology. Functional Ecology 2004; 18: 257-82.
		The theory based allometric exponent for functional characteristics such as CL is ¾. For convenience this is usually written as 0.75 but the theoretical value is derived from a ratio.

Slide 18	**Using Renal Function**
	
	• CL processes are assumed to be parallel hence
	$Cl_{total} = CL_{NR} + CL_R$
	
	• The covariate relationship has the same representation:
	
	`FSIZE=(WT/70)**0.75`
	`RF = CLCR/100 ; 100 mL/min 'normal' GFR`
	`GRPCL = (THETA(1) + THETA(2)* RF)*FSIZE`
	
	$Cl_{total} = CL_{NR} + CL_R$	

Slide 19

A Common Error Using CLCR

- CLCR predictions are scaled to the size of the individual e.g. Cockcroft & Gault uses a linear function of weight

  $CLCR_{CG}$ = (140-AGE)/SCR*WT/72*FSEX

- Using CLCR in this model for a drug only cleared by metabolism:

  RF = $CLCR_{CG}$/100 ; 100 mL/min 'normal' GFR

  GRPCL = THETA(1)* RF

  will appear to show an effect of CLCR because of the influence of WT used in the prediction of CLCR

- The solution is to use a size standardized CLCR

  $CLCR_{STD}$ = CLCRCG*70/WT

  RF = $CLCR_{STD}$/100 ; 100 mL/min 'normal' GFR for 70kg

  GRPCL = THETA(1)* RF

Note that CLCR should be calculated for a standard size person e.g. 70 kg otherwise RF will be change with weight as well as glomerular filtration rate. The Cockcroft & Gault formula prediction (CLCRCG) uses weight but can be easily used to obtain a size standardized CLCRSTD (see Mould et al 2002 for an example).
Mould DR, Holford NH, Schellens JH, Beijnen JH, Hutson PR, Rosing H, ten Bokkel Huinink WW, Rowinsky EK, Schiller JH, Russo M, Ross G. Population pharmacokinetic and adverse event analysis of topotecan in patients with solid tumors. Clinical Pharmacology & Therapeutics 2002; 71: 334-48.

Slide 20

Empirical Continuous Covariate Models 1

```
POPCL=THETA(1)

; Theory based allometry
FSIZE=(WT/70)**0.75

; Linear model
FAGE=1+THETA(2)*(AGE-20)

Or

; Exponential model
FAGE=EXP(THETA(2)*(AGE-20))

GRPCL=POPCL * FSIZE* FAGE

Note:
exp(x) = 1+x when x is small
```

POPCL is standardized on 70 and AGE=20

Linear model is simple but can lead to negative clearances if THETA(2) is negative and AGE is bigger than 20

Exponential model avoids negative clearance and is similar to linear model for small age effects

Whenever there is no plausible theory based model then empirical models are used. The effect of age on CL (in adults) has no good biological theory to provide a model.

Never consider age models without first accounting for size with an allometric model.

Slide 21

Empirical Continuous Covariate Models 2

```
POPCL=THETA(1)

; Theory based allometry
FSIZE=(WT/70)**0.75

; Exponential model
IF (AGE.GE.20) THEN ; adult
   FAGE=EXP(THETA(2)*(AGE-20))
ELSE
   FAGE=1
ENDIF

; maturation model
FMAT=1/(1+(PCA/TM50)**(-HILL))

GRPCL=POPCL * FSIZE * FAGE * FMAT
```

POPCL is standardized on AGE=20 for a mature adult

Exponential model for possible decrease of clearance with AGE

Sigmoid Emax model has biologically plausible extrapolation of 0 at post-conception age of 0 and 1 when fully mature

The time course of the maturation of clearance has to be described empirically but the 'ends' of the curve are known which means extrapolation to very young or very old age will be biologically plausible.

Slide 22

Empirical Discrete Covariate Models

```
POPCL = THETA(1)
FSIZE=(WT/70)**0.75
IF(SEX.EQ.1) THEN
   GRPCL=THETA(1)*FSIZE
ELSE
   GRPCL=THETA(2)*FSIZE
ENDIF
```
(1)

or

```
IF(SEX.EQ.1) THEN
   FSEX=1
ELSE
   FSEX=THETA(2)
ENDIF
GRPCL = POPCL*FSIZE*FSEX
```
(2)

```
GRPCL = POPCL*FSIZE*(1 + THETA(2)*SEX)
```
(3)

(1) Most basic model (probably not best for interpretation)

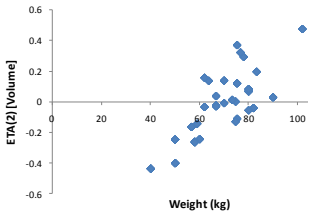
(2) Method 2 is more generic and can be applied for any number of discrete covariates

(2 & 3) Provide similar answers if SEX is coded as 1 and 0

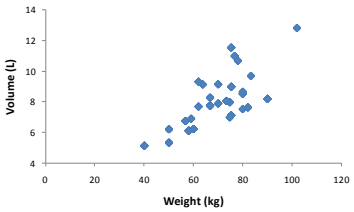
Slide 23

Assessing Potential Influence of Covariates with Base Model

V=THETA(2)*EXP(ETA(2))



ETA(2) vs WT



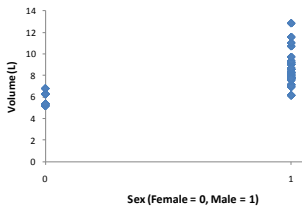
Volume vs WT

Note that without weight in the model there is seems to be a positive correlation between the ETA for volume and weight.

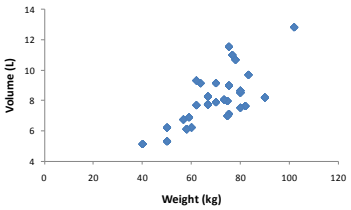
Slide 24

Assessing Potential Influence of Covariates with Base Model

Which covariate to choose?



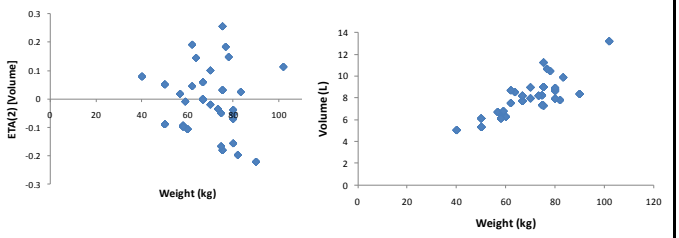
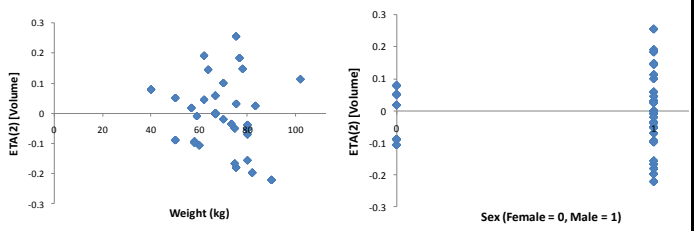
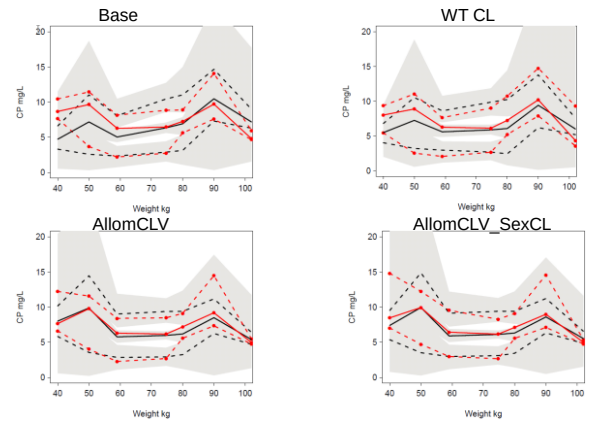
Volume vs Sex



Volume vs WT

Volume seems to be larger in males and also increases with weight. Which covariate should be considered?

Individual volume values are empirical Bayes estimates. These should not be relied upon to explore covariate relationships unless parameters are well estimated in each individual ('low shrinkage')

Slide 25	**Assessing Potential Influence of Covariates with Group Model** **V = THETA(2)*WT/70*EXP(ETA(2))**  ETA(2) vs WT Volume vs WT 	Note that with weight included as a covariate the relationship between weight and the individual volume appears even stronger (less scatter). The random effect for each individual does not seem to be a function of weight anymore.
Slide 26	**Assessing Potential Influence of Covariates with Group Model** "Put weight on -- lose interest in sex"  ETA(2) vs WT ETA(2) vs Sex 	With weight in the model the ETA for volume does not seem to be associated with sex any longer. This is biologically plausible because males are typically heavier than females.
Slide 27	**Visual Predictive Check**  	The VPC for the base model shows concs are underpredicted at low weights and perhaps overpredicted at high weights. Adding WT on CL (linear) improves the VPC a bit at low weights. Adding allometrically scaled WT on CL and V improves the fit a lot more. There is no convincing difference when SEX is added on CL.

Slide 28

Steps in Doing a Population Analysis

1) Search literature for previous population analyses
2) Write an analysis plan
3) Collect and clean data (never delete data points – just "comment them out")
4) Perform initial data summaries (plots etc)
5) Graph all covariates vs all covariates to find co-linearity
6) Prepare data for NONMEM input format
7) Initial model run – check and comment outliers
8) Model data
9) Evaluate Model
10) Write report

Slide 29

Analysis Plan

- An analysis plan is equivalent to a study protocol – except the term plan is not as definitive (i.e. it could be modified if necessary)
- Items to include:
  - Modelling tools
  - All statistical tests and tests of significance
  - Model structure and models to be considered (including covariates)
  - Models for random effects
  - Model evaluation techniques

Slide 30

Modeling Workflow

1. Run an Initial Model with No Estimation, check:
   - Model predictions versus observations
   - For outliers
2. Identify Base Model
   - Estimate residual error models (additive, proportion and combined error models)
   - Estimate structural model (e.g. 1 or 2 cpt)
   - Estimate random effects components
3. Identify Potential Covariate Relationships
   - Use your brain not posthoc values versus covariates
4. Include Covariates into Final Model

Be prepared to change the above order 2-4 because due to nonlinearity there is no single recipe…

Slide 31

Important Points for Covariates

- Standardize population parameters
- Use theory based models
- Never test for weight effects
- Only estimate allometric exponents with a sufficient design

Slide 32

A Standard Example

A Pharmacokinetic Standard for Babies and Adults

NICK HOLFORD,[1] YOUNG-A HEO,[1] BRIAN ANDERSON[2]

[1]Department of Pharmacology & Clinical Pharmacology, University of Auckland, New Zealand

[2]Department of Anaesthesiology, University of Auckland, New Zealand

Received 28 February 2013; revised 12 April 2013; accepted 12 April 2013

Published online 6 May 2013 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/jps.23574