

Mats Karlsson

Prof. Mats Karlsson, is Professor of Pharmacometrics at the Department of Pharmacy, Uppsala University, Uppsala, Sweden. He started the pharmacometrics research group at Uppsala University where more than 100 PhD students and postdocs have had pharmacometrics training. His research focus is methodological development and application of pharmacometrics to drug development and drug usage.



Decisions based on longitudinal analysis

Mats Karlsson, PhD Professor of Pharmacometrics

Department of Pharmacy Uppsala University Uppsala, Sweden





Guidance for Industry

Exposure-Response Relationships — Study Design, Data Analysis, and Regulatory Applications

U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER) Center for Biologics Evaluation and Research (CBER) April 2003 CP



Exposure – response dose individualization assessment

CS-4

Stroke or SEE by Baseline CrCL mITT On-Treatment

Protocol-defined subgroup analysis





Exposure - response



Analysis shown for "typical" patient with normal renal function: 64 years old, CrCL of 100 mL/min, 24.8% with prior stroke, 29.2% with baseline aspirin use, body weight of 95.3 kg



Guidance for Industry Population Pharmacokinetics

U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER) Center for Biologics Evaluation and Research (CBER)

> February 1999 CP 1



Pop PK in exposure matching

Clarification on Precision Criteria to Derive Sample Size When Designing Pediatric Pharmacokinetic Studies

Yaning Wang, PhD, Pravin R. Jadhav, PhD, Mallika Lala, PhD, and Jogarao V. Gobburu, PhD

Keywords: pediatric drug development; pharmacokinetics; regulatory requirement; precision Journal of Clinical Pharmacology, 2012;52:1601-1606 © 2012 The Author(s)



Workflow for pediatric studies





PKPD model from adults to children

Scale

Example: Comparison of scaling approaches for vancomycin (main elimination by glomerular filtration)[1,2]



[1] Parameter value from:

Anderson et al. "Vancomycin pharmacokinetics in preterm neonates and the prediction of adult clearance." Br J Clin Pharmacol 2007; 63 (1): 75-84

[2] Growth data from:

WHO Multicentre Growth Reference Study Group. "WHO Child Growth Standards based on length/height, weight and age". Acta Paediatr, Suppl. 2006, 450, 76-85.

de Onis M et al. "Development of a WHO growth reference for school-aged children and adolescents" Bull WHO, 2007;85:660-7.



Workflow for pediatric studies





Power

• Sample size needs to be chosen to fulfill precision criteria: *".. target a 95% CI within 60% and 140% of the geometric mean estimates of clearance and volume of distribution ... in each pediatric sub-group with at least 80% power." [1]*

Methodology:

- 1. Determine expected parameter standard errors (SE) for sample size N using CTS
- 2. Scale parameter SE and determine approximate sample size for intended precision
- 3. Verify approximate sample size calculation using CTS (asymp. nature of scaling, non-normality of estimates)

[1] Yaning Wang et al. "Clarification on precision criteria to derive sample size when designing pediatric pharmacokinetic studies." J Clin Pharmacol 2012;52:1601-1606



Workflow for pediatric studies





Motivating example for use of longitudinal modelling as primary analysis

- Proof-of-Concept study in gastroesophageal reflux
 - -Back-up compound
 - -Protocol for planned PoC already developed
 - Prior internal experience of PoC-study with same endpoint
 - •Data
 - Longitudinal NLME model
 - –Could NLME offer any advantage to planned design/analysis?

UPPSALA UNIVERSITET

Data from previous trial

Clinical symptom data in reflux disease





Two options in testing for a drug effect

• Based on a responder definition

–Maximally 1 symptom during last week

- -Test for differences in responder rates (RR)
- Based on a non-linear mixed effects model for longitudinal ordered categorical data (OC model)
 - -Test for a non-zero drug effect

Extensive diagnostics to assure sufficient quality of model on prior data e.g. visual predictive check (VPC)







Analysis alternatives for the PoC study Power to detect a clinically relevant effect





Citation: CPT: Pharmacometrics & Systems Pharmacology (2013) 2, e23; doi:10.1038/psp.2012.24 © 2013 ASCPT All rights reserved 2163-8306/12

www.nature.com/psp

ORIGINAL ARTICLE

Comparisons of Analysis Methods for Proof-of-Concept Trials

KE Karlsson¹, C Vong¹, M Bergstrand¹, EN Jonsson^{1,2} and MO Karlsson¹

Drug development struggles with high costs and time consuming processes. Hence, a need for new strategies has been accentuated by many stakeholders in drug development. This study proposes the use of pharmacometric models to rationalize drug development. Two simulated examples, within the therapeutic areas of acute stroke and type 2 diabetes, are utilized to compare a pharmacometric model-based analysis to a *t*-test with respect to study power of proof-of-concept (POC) trials. In all investigated examples and scenarios, the conventional statistical analysis resulted in several fold larger study sizes to achieve 80% power. For a scenario with a parallel design of one placebo group and one active dose arm, the difference between the conventional and pharmacometric approach was 4.3- and 8.4-fold, for the stroke and diabetes example, respectively. Although the model-based power depend on the model assumptions, in these scenarios, the pharmacometric model-based approach was demonstrated to permit drastic streamlining of POC trials.

CPT: Pharmacometrics & Systems Pharmacology (2013) 2, e23; doi:10.1038/psp.2012.24; advance online publication 16 January 2013



Pharmacometric vs. traditional approach to establish PoC

Stroke

•Traditional:

 –NIH stroke scale change from baseline at end of treatment (~90 days), LOCF for drop-out

-Two sided t-test (p<0.05) placebo vs. top dose level

• Model based:

The AAPS Journal (© 2010) DOI: 10.1208/s12248-010-9230-0

Research Article

Modeling Disease Progression in Acute Stroke Using Clinical Assessment Scales

Kristin E. Karlsson,^{1,5} Justin J. Wilkins,¹ Fredrik Jonsson,² Per-Henrik Zingmark,³ Mats O. Karlsson,¹ and E. Niclas Jonsson^{1,4}



Benefits of using all available data; NIHSS observations

Observations made at day 0, 7, 30 and 90





Data used in a traditional analysis; NIHSS observations





Data used in pharmacometric analysis; NIHSS observations



Day



PoC in Stroke



Pharmacometric power gain:

- Use longitudinal data
- Sensible handling of drop-out information



Design of a PoC study for an antidiabetic study





Information utilized with the traditional approach

Information utilized with the pharmacometric approach



Pharmacometric vs. traditional approach to establish PoC

Diabetes

• Model used in model-based analysis:

nature publishing group

ARTICLES

Models for Plasma Glucose, HbA1c, and Hemoglobin Interrelationships in Patients with Type 2 Diabetes Following Tesaglitazar Treatment

B Hamrén^{1,2}, E Björk³, M Sunzel⁴ and MO Karlsson²

¹Department of Medical Science, Clinical Pharmacology, AstraZeneca R&D Mölndal, Mölndal, Sweden; ²Department of Biopharmaceutical Sciences, Division of Pharmacokinetics and Drug Therapy, Uppsala University, Uppsala, Sweden; ³Clinical Cardiovascular, AstraZeneca R&D Wilmington, Wilmington, Delaware, USA; ⁴Department of Medical Science, Clinical Pharmacology, AstraZeneca R&D Wilmington, Wilmington, Delaware, USA; ⁴Department of Medical Science, Clinical Pharmacology, AstraZeneca R&D Wilmington, Wilmington, Delaware, USA. Correspondence: B Hamrén (bengt.hamren@astrazeneca.com)

Received 31 October 2007; accepted 2 January 2008; advance online publication 19 March 2008. doi:10.1038/clpt.2008.2

CLINICAL PHARMACOLOGY & THERAPEUTICS



FPG – HbA1c model



Figure 2. Schematic representation of the mechanism-based model for the FPG-HbA1c relationship,



VPC for FPG-HbA1c model FPG





VPC for FPG-HbA1c model HbA1c





PoC in Diabetes



Pharmacometric power gain:

- Use longitudinal data
- FPG and HbA1c simultaneously



Models for ADAS-Cog in Alzheimer's disease

Simulated data – power to detect a hypothetical drug effect



Reference: Ueckert et al. Pharm Res 31 (2014)



A comparison of five approaches to decision-making for a first clinical trial of efficacy

Simon Kirby^a* and Christy Chuang-Stein^b









Simulating total score using IRT model



Results for a placebo-controlled PoC trial in COPD using the EXACT score







Results for a placebo-controlled PoC trial in COPD using the EXACT score 90% CI

Approach	Mean (90%CI) difference in total score between arms FULL SCORE	Mean (90%CI) difference in total score between arms REDUCED SCORE	CI width FULL SCORE	CI width REDUCED SCORE	Relative sample size FULL SCORE	Relative sample size REDUCED SCORE
Standard – week 52	-3.34 (-8.08, 1.38)	-1.82 (-4.59, 0.94)	9.46	5.53	1	1
$IRT - different P_{MAX}$	-2.40 (-5.07, 0.29)	-1.21 (-2.77, 0.80)	5.36	3.57	3.1*	2.4*
IRT – different P _{MAX} & T _{PROG}	-2.49 (-6.26, 1.27)	-1.22 (-3.39, 1.00)	7.53	4.39	1.6	1.6

* This means that using the **standard approach** the sample size should be **3.1/2.4 times larger** to achieve the precision obtained with IRT model

Framework for M&S in regulatory review according to impact on regulatory decision





Manolis et al. CPT:PSP 2013:2 e31



Aspects to consider in model-based drug effect estimation

•1. Selection bias

 The model building process, where typically the best of many tested models is selected, may result in drug effect bias

•2. Model misspecification

- Models for longitudinal data are always misspecified to some degree.
 Consequences may include:
 - Inflated Type 1 error rate from misspecified placebo model
 - Loss of power from misspecified drug model
 - Bias in drug effect estimates from misspecified drug or placebo model



Model averaging



Model averaging in pharmacometrics

UNIVERSITET		
	GrossMark	
ORIGINAL PAPER		
Model selection and averaging of nonlin for robust phase III dose selection	near mixed-effect models	
Yasunori Aoki ^{1,2}		
	Research Article	
	Comparison of Model Averaging and Model Selection Analyzed by Nonlinear Mixed Effect Models	in Dose Finding Trials
	Simon Buatois, ^{1,2,3,5} Sebastian Ueckert, ⁴ Nicolas Frey, ¹ Sylvie Retout, ^{1,2} and	l France Mentré ³
	WILEY	



WHITE PAPER

Advanced Methods for Dose and Regimen Finding **During Drug Development: Summary of the EMA/EFPIA** Workshop on Dose Finding (London 4–5 December 2014)

FT Musuamba^{1,2,3*}, E Manolis^{1,4}, N Holford⁵, SYA Cheung⁶, LE Friberg⁷, K Ogungbenro⁸, M Posch⁹, JWT Yates⁶, S Berry¹⁰, N Thomas¹¹, S Corriol-Rohou⁶, B Bornkamp¹², F Bretz^{9,12}, AC Hooker⁷, PH Van der Graaf^{13,14}, JF Standing^{1,15}, J Hay^{1,16}, S Cole^{1,16}, V Gigante^{1,17}, K Karlsson^{1,18}, T Dumortier¹², N Benda^{1,19}, F Serone^{1,17}, S Das⁶, A Brochot²⁰, F Ehmann⁴, R Hemmings¹⁶ and I Skottheim Rusten^{1,21}

Inadequate dose selection for confirmatory trials is currently still one of the most challenging issues in drug development, as illustrated by high rates of late-stage attritions in clinical development and postmarketing commitments required by regulatory institutions. In an effort to shift the current paradigm in dose and regimen selection and highlight the availability and usefulness of well-established and regulatory-acceptable methods, the European Medicines Agency (EMA) in collaboration with the European Federation of Pharmaceutical Industries Association (EFPIA) hosted a multistakeholder workshop on dose finding (London 4–5 December 2014). Some methodologies that could constitute a toolkit for drug developers and regulators were presented. These methods are described in the present report: they include five advanced methods for data analysis

There was agreement at the workshop on the fact that selection of dose for phase III is an estimation problem and ons were also discussed during the workshop; however, mostly for should not be addressed via hypothesis testing.

s, quantitative systems pharmacology models, MCP-Mod, and model vization (Fisher information matrix (FIM)-based methods, clinical trial alue and limitations of these methods as well as challenges for their

implementation. Some applications in unerent merapeutic areas are also summarized, in line with the discussions at the workshop. There was agreement at the workshop on the fact that selection of dose for phase III is an estimation problem and should not be addressed via hypothesis testing. Dose selection for phase III trials should be informed by well-designed dosefinding studies; however, the specific choice of method(s) will depend on several aspects and it is not possible to recommend a generalized decision tree. There are many valuable methods available, the methods are not mutually exclusive, and they should be used in conjunction to ensure a scientifically rigorous understanding of the dosing rationale. CPT Pharmacometrics Syst. Pharmacol. (2017) 6, 418-429; doi:10.1002/psp4.12196; published online 0 Month 2017.



Model averaging for determining doseresponse from longitudinal data

	CrossMark		
ORIGINAL PAPER			
Model selection and averaging of nonlinear mixed-effect models for robust phase III dose selection Yasunori Aoki ^{1,2}			



Placebo Model

- Placebo Model + Linear
- Placebo Model + Log-linear
- Placebo Model + Emax
- Placebo Model + Sigmoidal

 $\begin{pmatrix} \theta_1, \theta_2, ..., \omega_1, \omega_2, ..., \sigma_1, \sigma_2 \\ (\theta_1, \theta_2, ..., \omega_1, \omega_2, ..., \sigma_1, \sigma_2) \\ (\theta_1, \theta_2, ..., \omega_1, \omega_2, ..., \sigma_1, \sigma_2) \\ \end{pmatrix}$

 $\begin{pmatrix} \theta_1, \theta_2, ..., \omega_1, \omega_2, ..., \sigma_1, \sigma_2 \\ (\theta_1, \theta_2, ..., \omega_1, \omega_2, ..., \sigma_1, \sigma_2) \\ (\theta_1, \theta_2, ..., \omega_1, \omega_2, ..., \sigma_1, \sigma_2) \\ \vdots$

Bootstrap $\begin{pmatrix} \theta_1, \theta_2, ..., \omega_1, \omega_2, ..., \sigma_1, \sigma_2 \end{pmatrix}$ $\begin{pmatrix} \theta_1, \theta_2, ..., \omega_1, \omega_2, ..., \sigma_1, \sigma_2 \end{pmatrix}$ $\begin{pmatrix} \theta_1, \theta_2, ..., \omega_1, \omega_2, ..., \sigma_1, \sigma_2 \end{pmatrix}$

 $\begin{array}{c} (\theta_1, \theta_2, ..., \omega_1, \omega_2, ..., \sigma_1, \sigma_2) \\ (\theta_1, \theta_2, ..., \omega_1, \omega_2, ..., \sigma_1, \sigma_2) \\ (\theta_1, \theta_2, ..., \omega_1, \omega_2, ..., \sigma_1, \sigma_2) \\ \vdots \end{array}$



+ PhIIb clinical trial data



Model weighting by goodness-of-fit



Effect

 $\times \frac{e^{-\text{AIC}_4/2}}{\sum_{i=1}^4 e^{-\text{AIC}_i/2}} \bullet \quad \text{Weighting scheme} \\ \text{proposed by Buckland et} \\ \text{al. 1997} \\ \text{Weighting scheme} \\ \text{weighting scheme} \\ \text{proposed by Buckland et} \\ \text{weighting scheme} \\ \text{weighting scheme} \\ \text{proposed by Buckland et} \\ \text{weighting scheme} \\$

This model averaging methodology combines both the parameter estimation uncertainty and model structure uncertainty to quantify overall uncertainty







Probability of Success based on Quantified Overall Uncertainty





Simulation Studies based on AZD 1981

	Frequencies of making the correct dose selection		
	Study Protocol (ANOVA + Averaged Effect)	Averaged Model Based	
Case 1 correct dose = 10mg	582	788	35% improvement
Case 2 correct dose = 40mg	361	592	64% improvement
Case 3 correct dose = 100mg	312	432	38% improvement
Case 4 correct dose = 400mg	402	519	29% improvement

Aoki et al. J Pharmacokinet Pharmacodyn 2017



Received: 16 April 2019 Revised: 14 December 2020 Accepted: 1 February 2021

DOI: 10.1002/sim.8913

RESEARCH ARTICLE

Statistics in Medicine WILEY

cLRT-Mod: An efficient methodology for pharmacometric model-based analysis of longitudinal phase II dose finding studies under model uncertainty

Simon Buatois^{1,2} | Sebastian Ueckert³ | Nicolas Frey² | Sylvie Retout² | France Mentré¹



MCP-MOD'	PMX
 Starting from a predefined set of dose- response candidate models: 	
I. MCP-step: Assessment of dose-response signal using contrast test on the best model (MS)	1. Model building using multiple LRT on nonlinear mixed effect models (MS)
2. MOD-step: Estimate the dose-response curve using either model selection (MS) or model averaging (MA)	2. Estimate the dose-response curve using the selected model
Advantages vs PMX	Advantages vs MCP-MOD
Models pre-specified	Longitudinal analysis of the data
Takes model uncertainty into account	3
Control the type l error	

Introduction

duction	MCP-MOD' Bes	st of PMX
Intro	 Starting from a predefined set of dc both w response candidate models: 	vorlds ?)
	I. MCP-step: Assessment of dose-response signal using contrast test on the best model (MS)	1. Model building using multiple LRT on nonlinear mixed effect models (MS)
	2. MOD-step: Estimate the dose-response curve using either model selection (MS) or model averaging (MA)	2 . Estimate the dose-response curve using the selected model
	Advantages vs PMX	Advantages vs MCP-MOD
	Models pre-specified	 Longitudinal analysis of the data
	 Takes model uncertainty into account Control the type I error 	3

MCP-N	10D'	cLRT-N	10D
	Predefined set of dos mod	e-response candidate lels:	
I. MCP-step: Assessment of dose-response signal using contrast test on the best model (MS)		I. cLRT-step: Assessme signal using a corrected-	ent of dose-response Likelihood Ratio Test ²
	2. MOD-step: Estima curve using either mo model aver	te the dose-response odel selection (MS) or aging (MA)	
[1] Bretz F . <i>et al</i> , Biometrics,	2005	[2] Dette H. et al, Biometrics, 20'	15

Methods

I. Corrected-LRT step:

Observed dataset:



[1] Aoki Y. et al, JPKPD, 2017[2] Buatois S. et al, AAPS, 2018

6



Assessments of a treatment effect with Individual Model Averaging (IMA)

	Base model (H0)	Full model (H1)	
STD	f(BASE, PLB)	f(BASE, PLB, DRUG * TRT)	
	Mixture 1: f (BASE, PLB) Mixture 2: f (BASE, PLB, DRUG)		
IMA	Mixture probability fixed to randomization: P(1) = 0.5 P(2) = 1 - P(1)	Mixture probability estimated: $P(1) = TRT^{*}(1 - \theta_{MIX}) + (1 - TRT)^{*} \theta_{MIX}$ $P(2) = 1 - P(1)$	

Chasseloup et al PAGE 28 (2019) Abstr 9149 [www.page-meeting.org/?abstract=9149]



Clinical data sets – placebo or natural history data

Likert pain score (categorical data) from 230 patients with neuropathic pain^{3,4} followed-up for 4 months

Seizure (count data) from 500 epileptic patients⁵ followed-up for 3 months

ADAS-Cog (continuous data): natural history of 800 patients with Alzheimer's disease⁶ followed-up for 36 months



Procedure to assess type 1 error

- For each data set (Pain, Seizures and ADAS-Cog), placebo patients were randomized 1:1 to TRT=0 ("placebo") or TRT=1 ("active")
- This creates a data set similar to a two-arm trial, where the "active" treatment has zero effect
- The randomization was repeated to generate 1000 data sets for each disease

ADAS-Cog data – STD and IMA





ADAS-Cog – drug effect estimates

UNIVERSITET Adas–cog data

UPPSALA





Type 1 error rates – STD and IMA Pain, Seizures & ADAS-Cog





Drug models Type I error 0-3.72 6.55-10 30-60 80-100 3.73-6.54 10-30 60-80



Summary IMA vs STD approach

- •Type 1 error rates:
- typically inflated for standard approach
- controlled for individual model averaging (IMA)
- Drug effect estimates
- typically biased for standard approach
- unbiased for individual model averaging (IMA)



Rare diseases: Example of single-subject customized antisense oligonucleotide treatment

BRIEF REPORT

The NEW ENGLAND JOURNAL of MEDICINE

Patient-Customized Oligonucleotide Therapy for a Rare Genetic Disease

V

e

а



- Genetic diagnosis
- AON design
- Tests in fibroblasts
- Tox studies in rats
 - First treatment

Kim et al, 2019, NEJM



Milasen treatment: first year



Kim et al, 2019, NEJM



Response from FDA

Drug Regulation in the Era of Individualized Therapies

Janet Woodcock, M.D., and Peter Marks, M.D., Ph.D.

Kim et al., in a report now published in the Jour- known as CLN7), and a previously undescribed nal,1 describe the discovery, development, and insertion of a retrotransposon was present in the administration of an antisense oligonucleotide other copy. Retrotransposons are stretches of (ASO) therapy specifically designed for a single DNA that are sometimes described as mobile patient with CLN7 neuronal ceroid lipofuscino- elements; thousands are present in the human sis (a form of Batten's disease), a fatal genetic genome, and some are capable of moving to a neurodegenerative disorder.² In this patient, a new location — such as the middle of a gene known pathogenic point mutation was found to — through a "copy and paste" mechanism. The be present in one copy of the gene MFSD8 (also authors showed that the retrotransposon inser-

tion in this patient led to missplicing of the

MFSD8 messenger RNA (mRNA) and probably to premature translational termination. The authors devised candidate ASOs that were intended to "correct" the missplicing of the mRNA and selected a candidate ASO that, in cultured patient fibroblasts, resulted in an increase in the ratio of normal to mutant mRNA. Evaluation of lysosomal function in vitro showed improvements in the presence of the ASO. After an abbreviated toxicologic evaluation and after obtaining authorization from the Food and Drug Administration (FDA) and expedited institutional review board approval,

could ultimately be treated affect the decisionmaking process?

In addition, how will efficacy be evaluated? At the very least, during the time needed to discover and develop an intervention, quantifiable, objective measures of the patient's disease status should be identified and tracked, since, in an N-of-one experiment, evaluation of disease trends before and after treatment will usually be the primary method of assessing effectiveness. In this regard, there is precedent for the application of new efficacy measures to the study of small numbers of patients.5



Concluding remarks

- Longitudinal data analysis using pharmacometric (NLME) models are already an important tool in drug development
- Such analyses use available data very efficiently and are for that reason attractive
- To date, their use in primary analyses has been limited, but this may change as
 - -the insight in their advantages is increasing
 - -the methodology for their use is maturing
 - -situations where good alternatives are lacking is increasing