Is precision dosing worth the effort?

An overwhelming number of medicinal products are marketed with the same fixed dose for every patient. This is suboptimal. Each patient has distinctly different pharmacokinetics and pharmacodynamics, and should therefore receive an individualized dose. We can even precisely calculate this dose by observing the individual response and adapting based on a non-linear mixed effects statistical model. Unfortunately, this requires a lot of effort: the medical practitioner should sample blood, assay drug concentration, and use specialized computer software to calculate the individual dose.

Is this worth the effort? For all drugs, or for some?
In this doctoral research, we simulate the impact of precision dosing on virtual patients. We showed that precision dosing of infliximab will not help to reduce costs in the induction treatment of ulcerative colitis. For tacrolimus, an immunosuppressor used to mitigate rejectin in kidney transplant patients, we showed that precision dosing will lead to safe and effective tacrolimus blood concentrations faster and more often. It was only a small step to go from virtual patients to real patients in clinic, through the development of a bedside dose recommendation app. A prospective clinical trial shows promising interim results.

We hope this approach may be used to evaluate precision dosing of other drugs, and effectively develop solutions where precision dosing is worth the effort.

KU Leuven, September 2022
Doctoral thesis in Pharmaceutical Sciences

Ruben FAELENS

Computer-assisted dosing recommendation framework, feasibility study and proof of concept implementation for tacrolimus

Ruben FAELENS

Computer-assisted dosing recommendation framework, feasibility study and proof of concept implementation for tacrolimus

Paving the road towards model-informed precision dosing

KU Leuven
Biomedical Sciences Group
Faculty of Pharmaceutical Sciences
Department of Pharmaceutical and Pharmacological Sciences
Drug Delivery and Disposition

KU LEUVEN

DOCTORAL SCHOOL
BIOMEDICAL SCIENCES

Computer-assisted dosing recommendation framework, feasibility study and proof of concept implementation for tacrolimus

Ruben FAELENS

Promoter:	Prof. dr. Pieter Annaert
Co-promoter:	Prof. Dr. Dirk Kuypers
	dr. Daniel Röshammar
Chair:	Prof. dr. Isabelle Huys
Jury members:	Prof. Dr. Minne Casteels
	Prof. dr. Geert Verbeke
	dr. Philippe Jacqmin
	Prof. dr. Sebastian Wicha

Dissertation presented in partial fulfilment of the requirements for the degree of Doctor in Pharmaceutical Sciences

Leuven, September 7th 2022

Word of thanks

This project has been a long and arduous journey in many aspects. Pharmacometrics and precision dosing is an interesting field with applications close to patients. It is also encumbered by business interests, legislative concerns, politics, and deep-rooted convictions.

I wish to thank my first promotor Dr. Thomas Bouillon. He believed in my previous experience, and was able to leverage and challenge this experience to create a working and stable software in (comparatively) little time. He further managed to identify pressing needs in the hospital and on-board key players in the project. Prof. Apr. Pieter Annaert supported the project from the beginning, and I am truly grateful he continued the promotorship in the last years of my PhD. Pieter not only taught me how to work efficiently, he also stressed the importance of work-life balance, ensuring a durable career in the face of adversity. My colleagues at the department of Drug Delivery and Disposition offered an interesting change of perspective. Thank you Neel, Pieter, Nina, Julia and Miao-Chan for being warm colleagues, including me on projects, and accepting me in the group. Thank you Matthias for sharing your insights in why tacrolimus is such a variable compound. I also connected with fellow students at Clinical Pharmacology (researching the impact and cost of medicines), who offered yet another view on drug development. The team of Erwin Dreesen, Ann Gils, Paul Declerk, Zhigang Wang, Wannee Kantasiripitak and Isabel Spriet were a valuable sparring partner. Thank you. I further want to thank the Physics department of KULeuven, as they provided me with rewarding teaching opportunities within my competence level.

This project was only possible with the greater support of all people and departments it touched. Thank you Prof. Dr. Dirk Kuypers and team for the clinical drive to make this project a reality. Thank you Maxine for hitting the ground running and picking up the monitoring and support tasks for the clinical study. I leave the project in your capable hands. Thank you to the Biostats department for supporting the clinical trial. Thank you Dr. Van Hove for diligently assembling the valuable datasets that brought us this far. Thank you Bart, Egon and Thomas for the technical support from the KWS side. I also want to thank the wider research team at the university, and prof. Myriam Baes specifically, for their support in ensuring continued funding. In

this respect, I thank the FWO for believing in this project and sharing our vision.

Pharmacometrics is a special community. It connects people from different disciplines. This collaborative, open, supporting spirit was created by amazingly passionate individuals, too many to list them all here. I especially want to thank my ex-colleagues at Exprimo -Philippe Jacqmin, Per Olsson and Justin Wilkins in particular- for believing in my potential and spending countless hours to teach me. I am humbled by their support. Thank you to all of the members of the EU Mon4Strat consortium, in particular prof. Paul Tulkens, prof. Dr. Jean Chastre and the colleagues at UCL Liège for not dismissing my youthful enthousiasm but nurturing it. You supported my first exploration of precision dosing of meropenem and allowed me to grasp the broader context and real-life challenges. The Mon4strat project was a great first step, and I am proud to have been a part of it.

When I joined the pharmacometrics community, I was warmly welcomed. Independent of affiliation, we strive to support and develop new talent. The awesome students Louis Sandra and Mirthe Vincken allowed this learning cycle to continue; thank you for allowing me to teach you. As R. Heinlein said: “When one teaches, two learn.”. A warm thanks to all participants of the DDMoRe project, as they showed how difficult software development can be in the pharmacometrics space.

Thank you to Daniel Röshammar, my ex-manager and co-promotor, for all your hard work in guiding me towards growth as a pharmacometrician. You both supported me as a consultant at Exprimo, and in my search for career growth as a PhD student. Thank you Nicolas and Quentin for keeping the warm contact independent of our employers, and for sharing my passion on this project and the domain of MIPD.

Thank you to all my friends and family. You brought me perspective and support, even when I had very little to offer apart from a head full of problems. You offered a safe haven openly and freely. Thank you. I am looking forward to fulfilling my role as a godfather of Thijs and Jonas with a little more focus.

In theory, a PhD project requires full and dedicated focus. In practice, this is often challenged, and for the better. The building project of a full street with 33 houses consumed a part of my time and brainpower; I am grateful for the life lessons it brought. Thank you everyone who supported our struggle.

During my childhood, scientific inquiry was supported and cheered on. My dear parents, thank you for inspiring me to aim high, both professionally and personally. Thank you Femke and Wouter for always listening to my worries and to tolerate my big-brotherly advice. Uncle Charles, thank you for instilling a passion for science in me; I hope to carry on the torch.

Midst all work-related challenges, my wife Astrid and children Linde and Warre were always my first priority. Their infinite kindness to offer a place in their heart -even for a stressed and grumpy dad- makes me smile and tear up. Even when big dreams turned out difficult, our beautiful family is my dream come true.

Thank you all.

Table of Contents

Included only in the digital PDF version

Glossary

2LL 2-log likelihood
AB Adams-bashforth
ADME Absorption, Distribution, Metabolization and Excretion
ALT Alanine Aminotransferase
API Application programmer interface
ATI Antibodies to infliximab
AUC Area under the curve
BDF Backwards differential formula
BFGS Broyden-Fletcher-Goldfarb-Shanno algorithm
BSD Berkeley Software Distribution
BSV Between-subject variability
CAUC Cumulative AUC
CCT Concentration-controlled trial
CE Conformité Européenne
CF Cystic fibrosis
CI Confidence interval
CI Continuous integration
CL Clearance
CRP C-reactive Protein
CSV Comma-seperated values
CSV Computer systems validation
CU Clinical utility
CV% Coefficient of variation
CWRES Conditionally-weighted residual
CYP Cytochrome P450
DDI Drug-drug interaction
DV Dependent variable
EBE Empirical Bayesian Estimation
ECDF Empirical cumulative distribution function
EHR Electronic health record
EMA/EMEA European Medicines Agency
EUPL EU Public License
FDA Federal Drug Agency
FFM Fat-free mass
FIH First in human
FO First-Order estimation
FOCE First-Order Conditional estimation
FOCEI First-Order Conditional estimation with interaction
GDPR General Data Protection Regulation
GLS General linear system
GNU GNU's Not Unix
GoF Goodness of Fit
GPL GNU General Public License
HIPAA Health insurance portability and accountability act
HTML Hypertext markup language
IBD Inflammatory bowel disease
IIV Inter-individual variability
IM Intramuscular
IOV Inter-occasion variability

IPRED Individual predicted value
IQ Installation qualification
IQR Inter-quartile range
ISoP International Society of Pharmacometrics
IT Information Technology
IV Intravenous
IWRES Individual weighted residual
JSON JavaScript Object Notation
KA Absorption rate
KS-test Kolmogorov-Smirnov test
KWS Klinisch werkstation
LC-MS/MS Liquid chromatography-mass spectrometry
LOCF Last Observation Carry Forward
LSODA Livermore Solver for Ordinary Differential equations with Automatic stiff/non-stiff selection
MAP Maximum a posteriori estimate
MCMC Markov chain monte carlo
MDE Minimal detectable effect
MID3 Model-informed drug discovery and development
MIDD Model-informed drug development
MIPD Model-informed precision dosing
MIT Massachusetts Institute of Technology
MMF Mycophenolate Mofetil
MMRM Mixed model repeated measures
MPC/MIPD Model-predictive control for MIPD
MPPE Mean percentage prediction error
MSE Mean squared error
NLME Nonlinear mixed effects
NOCB Next Observation Carry Backward
NONMEM Nonlinear Mixed Effects Modeling program
NPAG Non-parametric adaptive grid
NPDE Normalized prediction distribution error
ODE Ordinary differential equations
OFV Objective function value
OQ Operational qualification
PBPK Physiology-based Pharmacokinetics
pcVPC prediction-corrected visual predictive check
PD Pharmacodynamics
pEI Probability of Endoscopic Improvement
PI Prediction interval
PK Pharmacokinetics
popPK/PD Population pharmacokinetics/pharmacodynamics
PoSS Probability of Study Success
PPI Proton pump inhibitor
PQ Performance qualification
PRED Population-predicted value
PsN Perl speaks nonmem
PTA Probability of Target Attainment

QALY Quality-adjusted life years
R R statistical programming language
RCT Randomized controlled trial
RE Residual error
RIZIV Rijksinstituut voor ziekte- en invaliditeitsverzekering
RK Runge-kutta
RMSE Root mean squared error
RSE Relative standard error
RxODE R package for simulation of ordinary differential equations
SAEM Stochastic approximation and expectation maximization
SaMD Software as a Medical Device
SC Subcutaneous
SCM Stepwise covariate modeling
SDE Stochastic differential equations
SLD Sum of largest tumor diameter
SoC Standard of care
T50 Time of 50% of effect
TAC Tacrolimus
TCI Target concentration intervention
TDM Therapeutic drug monitoring
TNFalpha Tumor necrosis factor alpha
TTE Time To Event
TVx Typical value of x
UC Ulcerative Colitis
UZ Leuven Universitair Ziekenhuis Leuven
V Distribution volume
VPC Visual predictive check
WSV Within-subject variability

Popular Summary

Everyone is different. Some people are small, others are large. We all need different amounts of food. We can handle different pain levels. While everyone is different, most common medicines use the same dose for everyone. In theory, *individualizing* the dose for every patient should work better. As an example, we do this for the painkiller paracetamol in children: we give 15mg for every kg of body weight.

For some medicines, individualizing the dose is more difficult. Some patients may need 10 times more medicine than others, and there is no way of knowing this up front. For those medicines, we start with a standard dose, and then measure what happens to the patient. If the dose was too high or too low, we can adapt and try again, until we eventually find the right dose for that patient. Using a computer is the most accurate way to calculate the optimal dose. We call this *model-informed precision dosing*. We only do this for medicines where finding the right dose is really important. For most medicines, the benefit does not outweigh the extra costs and effort involved.

In this PhD thesis, we predicted how well a computer can find the right dose by running computer simulations. We did this for two medicines: infliximab and tacrolimus. Infliximab is given to patients with a disease where the immune system attacks the bowels. This can cause bloody diarrhea and malnutrition. Infliximab suppresses a small part of the overactive immune system that causes this disease. It is an expensive medicine however, and scientists believed that individualizing the dose could treat more people with the same amount of medicine. We predicted what would happen, and found that individualizing the dose actually uses more medicine per patient, without more people getting better.

The second medicine is tacrolimus. Patients who receive a new kidney from a donor need this drug to suppress their immune system, because the body wants to attack the foreign donor kidney. Tacrolimus is toxic to kidneys in high amounts, but too low amounts will not suppress the immune system enough and will also cause damage. Doctors measure the amount of tacrolimus in a patient's blood every day, and use this to adapt the dose. We predicted whether a computer could adapt the dose better. We found that the computer did this better and faster. We also predicted that you could prove this in a clinical trial with 200 patients. We built software to adapt the dose

for real patients, and this software is used in the university hospital of Leuven.

We also looked at what happens when tacrolimus and voriconazole/posaconazole are used together. Voriconazole and posaconazole are medicines that help kill germs. They are given to patients with transplanted organs. These patients need help from medicines, because their immune system is suppressed. Because tacrolimus and voriconazole/posaconazole are both removed from the body by the liver, they are both removed more slowly when these medicines are given together. We should then give less tacrolimus every day: doctors need to adapt the dose. We found that doctors should divide the tacrolimus dose by 3. We also found that this is different per patient, but there is no way of knowing this up front. Computers will not help.

We built software to make these predictions, and we called it *tdmore*. We made it easy to adapt, so other scientists can also predict if the computer can help dose other medicines. Unfortunately, we could not share the software that was designed for doctors, because the european union medical device law forbids sharing this without (expensive) validation tests.

This work has improved our knowledge about model-informed precision dosing. We built software to predict the effect of precision dosing, and showed how to calculate if precision dosing will benefit patients. If not, this is a strong argument for using the same dose for everyone. If yes, scientists can easily build precision dosing software for use in real patients.

Populairwetenschappelijke samenvatting

Iedereen is anders. Sommige mensen zijn groot, anderen zijn klein. Dik, dun, een veelvraat of een kleine eter, kleinzerig of ongevoelig. Hoewel iedereen verschilt, gebruiken we meestal dezelfde dosis van een geneesmiddel voor iedereen. In theorie zou het *individualiseren* van de dosis beter moeten werken. We doen dit al voor kinderen: voor de pijnstiller paracetamol geven we bijvoorbeeld 15mg per kg lichaamsgewicht.

Voor sommige geneesmiddelen is dat *individualiseren* moeilijk. Sommige patienten hebben 10 keer meer nodig dan anderen, en er is geen enkele manier om dit op voorhand te bepalen. Voor deze geneesmiddelen beginnen we vaak met dezelfde dosis voor iedereen. We meten bij de patiënt en passen de dosis steeds aan tot het goed is. Een computer moet dit in principe het meest nauwkeurig kunnen voorspellen met wiskundige formules. Dit noemen we *model-informed precision dosing*. Het is enkel de moeite om dit te doen voor geneesmiddelen waar de juiste dosis heel belangrijk is. In de meeste andere gevallen weegt het voordeel niet op tegen de extra moeite en kosten.

In dit doctoraat voorspellen we hoe goed een computer de juiste dosis kan vinden. Dit doen we aan de hand van computersimulaties op virtuele patiënten. We deden dit voor twee geneesmiddelen: infliximab en tacrolimus. Infliximab wordt gebruikt om patiënten met prikkelbaar darmsyndroom te behandelen. In deze patiënten valt het immuunsysteem de darmen aan, en dit kan voor bloederige stoelgang en ondervoeding zorgen. Infliximab onderdrukt een kleine schakel in het overactieve immuunsysteem en kan zo de symptomen verbeteren. Infliximab is duur, en dokters dachten dat je met precisiedosering het geneesmiddel efficienter gebruikt. Zo kan je patiënten beter behandelen met hetzelfde geld. Wij voorspelden wat er zou gebeuren, en vonden dat precisiedosering eigenlijk meer geneesmiddel gebruikt per patiënt, zonder dat de patiënt hier beter van wordt.

We onderzochten ook het geneesmiddel tacrolimus. Nier-patiënten die een donor-orgaan krijgen getransplanteerd hebben dit medicijn nodig om hun immuunsysteem te onderdrukken. Het lichaam wil immers de

lichaamsvreemde donornier aanvallen. Om goed te werken, moet er voldoende tacrolimus gegeven worden, maar het geneesmiddel is ook giftig voor de nieren in te hoge dosissen. Dokters meten daarom dagelijks de concentratie tacrolimus in het bloed, en passen de dosis naargelang aan. We voorspelden dat een computer dit beter en sneller kan doen. We toonden ook aan dat een klinische studie dit kan bewijzen met 200 patiënten. Daarom bouwden we ook software om de dosis voor echte patiënten aan te passen. Deze software wordt momenteel gebruikt in het universitair ziekenhuis van Leuven.

We keken ook wat er gebeurt als tacrolimus en voriconazole/posaconazole samen toegediend worden. Voriconazole en posaconazole zijn twee geneesmiddelen die schimmels doden. Omdat transplantatiepatiënten een onderdrukt immuunsysteem hebben, hebben zij deze geneesmiddelen nodig om schimmelinfecties te voorkomen en te behandelen. Beide geneesmiddelen worden door de lever afgebroken, en deze afbraak gebeurt dus trager als beide samen worden toegediend. Gezien de afbraak trager verloopt, moeten we dus ook minder tacrolimus geven. We vonden dat dokters de dosis tacrolimus moeten delen door 3 wanneer ze met voriconazole of posaconazole starten. We vonden ook dat de optimale dosisaanpassing verschilt per patiënt, maar er is geen enkele manier om dit op voorhand te meten. Model-informed precision dosing zal hier dus niet helpen.

De software die we bouwden om deze voorspellingen te doen, heet *tdmore*. We maakten de software gratis en vrij beschikbaar. We maakten het ook makkelijk voor andere onderzoekers om andere geneesmiddelen te evalueren. Jammer genoeg konden we de dokter-software niet delen, omdat de Europese Unie dure officiële testen verplicht.

Dit onderzoek heeft onze kennis over model-informed precision dosing verbetert. We bouwden software om het effect van precision dosing te voorspellen, en toonden hoe je dit kan vertalen naar een voordeel voor patiënten. Onderzoekers kunnen eenvoudig evalueren of precisiedosering de moeite waard is. Zo ja, dan kan ook relatief eenvoudig software gebouwd worden zodat dokters ook de dosis in echte patiënten kunnen aanpassen.

Summary

An overwhelming number of medicinal products are marketed with the same fixed dose for every patient. Adapting the dose for each patient should result in superior efficacy and safety, at least in theory. The gold standard of dosing individualization is model-informed precision dosing (MIPD). An extensive dataset of many patients is used to identify a *population pharmacokinetic/pharmacodynamic model* (popPK/PD). This is composed of a mathematical model predicting outcomes (drug concentration, biomarkers, or clinical outcome) over time, the variability of parameters for that model between individuals, and any predictive covariates for these individually variable parameters. By then using observations of an individual patient, the model parameter values for that individual patient can be estimated. These parameter values are subsequently used to accurately predict future outcomes for a candidate dose, and select the most optimal future dose: model-informed precision dosing.

This thesis aims to pave the road towards MIPD by exploring two key aspects. First, we show *how to predict the effect of precision dosing in silico*. Similar to how model-informed drug development has rationalized the drug development process, quantifying the effect of MIPD allows us to make well-informed choices, optimize investment into high-value opportunities for clinical improvement, develop better models, design better dosing strategies, and design better trials to show benefit. Second, we *simplify building MIPD software tools through reusable software*. Such a reusable software framework and accompanying scientific methodology reduces MIPD implementation time and cost. To show these goals are achieved, we *apply this methodology to clinical use cases*: infliximab induction therapy for ulcerative colitis patients, and tacrolimus immunosuppressive therapy for kidney transplant recipients.

We first developed the *tdmore* software package, integrating pharmacometric models with individual parameter estimation and future dose optimization. The mathematical routines are accompanied by debugging tools, likelihood profile visualization, and population simulation. This software package and simulation methodology was applied to investigate precision dosing of infliximab induction therapy in ulcerative colitis. In silico, a gradual reduction in outcome variability was predicted when moving from fixed dosing to covariate-based MIPD and concentration-based MIPD.

Surprisingly, average mean dose per patient was predicted to increase, without an associated improvement to mean outcome. This important negative case-study predicted that precision dosing, contrary to popular belief, may not be appropriate for infliximab induction therapy, at least in the proposed implementation. This software and simulation methodology was also applied to tacrolimus dosing for kidney transplant recipients. We showed how early simulation of predictive performance can inform modeling decisions, leading us to discard covariates in favor of the base model, as well as implementing a new estimation technique to account for parameter drift. Population simulation of MIPD showed a clinically relevant improvement in probability of target attainment, speed of target attainment and -for patients not in target- distance to target window. Clinical trial simulation informed the clinical team to expand enrollment from 100 to 200 patients, reducing the probability of an expensive but ultimately inconclusive clinical study. Both use cases demonstrate the overarching goal of rational MIPD development. Based on these simulations, a precision dosing tool was developed. First, general-purpose user interface components were developed. These were combined with an automated exchange of patient data with the electronic patient record database, and extensive business rules for automated conversion of clinical data to pharmacometric data. This system is undergoing clinical trial testing at Leuven University Hospitals since April 2021.

In conclusion, this thesis has advanced the domain of MIPD. To the best of our knowledge, this is the first scientific work proposing a clear roadmap to perform informed development of MIPD: we showed how to develop a model fit-for-purpose, how to simulate whether MIPD would outperform standard of care, and which clinical trial can demonstrate this. From an engineering point of view, we greatly simplified the transformation from pharmacometric model to precision dosing tool. Finally, we showed clinical results: our work allowed future infliximab MIPD efforts to refocus their aims, and showed improved tacrolimus target attainment in silico. We developed a tacrolimus precision dosing tool in a cost-effective manner, and designed a prospective randomized clinical trial which is currently ongoing at University Hospitals Leuven.

Beknopte samenvatting

Een overweldigend aantal geneesmiddelen wordt op de markt gebracht met dezelfde dosis voor elke patiënt. In theorie zou het individualiseren van de dosis moeten resulteren in een betere werkzaamheid en minder bijwerkingen. De gouden standaard voor dit individualiseren is model-informed precision dosing (MIPD). Een dataset verzameld bij veel patiënten wordt gebruikt om een populatie farmacokinetisch/farmacodynamisch model (popPK/PD) te bouwen. Dit bestaat uit een wiskundig model dat de uitkomsten (drugsconcentratie, biomarkers of klinische uitkomst) in de tijd voorspelt, een statistisch model dat de variabiliteit van model-parameters tussen individuen beschrijft, en eventuele voorspellende covariaten voor deze individueel variabele parameters. Door vervolgens uitkomsten bij een individuele patiënt te meten, kunnen de modelparameterwaarden voor die patiënt worden geschat. Deze parameterwaarden worden vervolgens gebruikt om toekomstige uitkomsten voor een kandidaatdosis nauwkeurig te voorspellen en de meest optimale toekomstige dosis te selecteren: model-informed precision dosing.

Dit proefschrift wil de weg naar MIPD effenen door twee belangrijke aspecten te onderzoeken. Eerst laten we zien *hoe het effect van precisiedosering in silico* kan worden voorspeld. Net zoals modelgeïnformeerde medicijnontwikkeling het ontwikkelingsproces heeft verbeterd, laat het voorspellen van MIPD ons toe om beter geïnformeerde keuzes te maken, te investeren in gebieden waar MIPD wel degelijk tot verbetering zal leiden, om betere modellen te ontwikkelen, betere doseringsstrategieën te ontwerpen, en om betere studies te ontwerpen die dit voordeel aantonen. Om het bouwen van MIPD-softwaretools te vereenvoudigen, *ontwikkelden we herbruikbare software*. Een dergelijk herbruikbaar softwareraamwerk en bijbehorende wetenschappelijke methodologie vermindert de implementatietijd en -kosten van MIPD. Om aan te tonen dat deze doelen worden bereikt, *passen we deze methodologie toe op klinische usecases*: de opstart van infliximab-therapie voor patiënten met colitis ulcerosa, en tacrolimus immunosuppressieve therapie voor niertransplantatiepatiënten.

We hebben eerst het *tdmore*-softwarepakket ontwikkeld, waarin farmacometrische modellen worden geïntegreerd met routines voor individuele parameterschatting en toekomstige dosisoptimalisatie. Die

wiskundige routines worden vergezeld van grafische weergave van de waarschijnlijkheidscurve, en routines voor populatiesimulatie. Dit softwarepakket en deze simulatiemethodologie werden toegepast op infliximab-inductietherapie bij colitis ulcerosa. Er werd een vermindering van de variabiliteit in het klinisch resultaat aangetoond indien men van een vaste dosering naar covariabele-gebaseerde MIPD of concentratie-gebaseerde MIPD overschakelt. Verrassend genoeg nam de gemiddelde dosis per patiënt toe, zonder verbetering in het globaal klinisch resultaat. Deze belangrijke negatieve casus toonde aan dat een precisiedosering, in tegenstelling tot wat vaak wordt gedacht, mogelijk niet geschikt is voor infliximab-inductietherapie, althans in de voorgestelde modaliteit. We onderzochten ook de dosering van tacrolimus voor niertransplantatiepatiënten. We toonden aan hoe het kwantificeren van voorspellend vermogen van een model tot betere modelleringsbeslissingen leidt. Zo konden we covariaten verwerpen en een eenvoudig model behouden. Ook implementeerden we een nieuwe parameter-schattingstechniek om rekening te houden met tijdgebonden variabiliteit. Populatiesimulatie van MIPD toonde een klinisch relevante verbetering aan. Meer bepaald verhoogde de kans om concentraties in het therapeutisch venster te hebben, werd het therapeutisch venster sneller bereikt, en -voor patiënten die buiten het therapeutisch venster lagen- verkleinde de afstand tot het venster. Simulatie van de klinische studie gaf indicatie om rekrutering van 100 naar 200 patiënten op te trekken, waardoor een goedkopere maar nutteloze klinische studie werd vermeden. Beide use cases tonen het overkoepelende doel van rationele MIPD-ontwikkeling aan. Op basis van voorgaande simulaties werd een precisiedoseringstool voor tacrolimus ontwikkeld. Eerst werden generieke gebruikersinterfacecomponenten gebouwd, die werden gecombineerd met een geautomatiseerde uitwisseling van patiëntgegevens met de elektronische patiëntendossierdatabase en uitgebreide bedrijfsregels voor geautomatiseerde conversie van klinische gegevens naar farmacometrische gegevens. Dit systeem wordt sinds april 2021 klinisch getest in UZ Leuven.

We kunnen besluiten dat dit proefschrift heeft bijgedragen tot het domein van MIPD. Vanuit wetenschappelijk oogpunt is dit het eerste werk dat een duidelijke roadmap heeft opgesteld om een rationele ontwikkeling van MIPD uit te voeren: hoe kunnen we een model ontwikkelen dat geschikt is voor het doel, hoe simuleren we of MIPD beter zou presteren dan de standaardzorg, en welke klinische proef kan dit helder aantonen? Vanuit technisch oogpunt vereenvoudigden we de omzetting van farmacometrisch model naar nauwkeurig doseerinstrument. Ten slotte lieten we klinische resultaten zien:

toekomstige infliximab MIPD-inspanningen zullen op relevante doelen focussen, en MIPD verbeterde tacrolimus-dosering in silico. We ontwikkelden software voor precision dosing van tacrolimus op een kostenefficiënte manier, en ontwierpen een klinische studie die momenteel loopt in UZ Leuven.

Chapter 1 Introduction to pharmacometrics and precision dosing

"Daddy, why do you eat more than me?" — Linde Faelens

At dinner, my daughter asked a fundamental question. The answer seems straightforward: "I am bigger, so I need more energy to get through the day." Why then does over 90%[1] of drug labels recommend a single dosing regimen for the adult population?

In this chapter, we will introduce the concept, history and need for precision dosing from both a theoretical and practical point of view. Finally, we will scrutinize the state-of-the-art in the scientific domain investigated by this PhD thesis and formulate clear goals and scientific questions to be answered.

1.1. The road to safe and effective drugs

Fundamental biological research tries to describe biological processes in the human body, some of which are involved in a certain disease. When such a process is identified, any substep in that process is a potential target for interaction or modification, potentially treating the disease. Molecules are designed and tested for interaction with that target, and for unwanted off-target interactions. The goal of this drug discovery process is to find a molecule that has high target engagement, while limiting off-target engagement.

This can be quantified into an $EC50$ metric; the concentration at which a 50% effect is measured at the target site. Unwanted engagement can also be quantified into an $EC50$. The ratio of wanted to unwanted engagement is an indication for the margin of safety for the drug. Of course, a drug should not

[1] Wang et al., "A Systematic Assessment of US Food and Drug Administration Dosing Recommendations For Drug Development Programs Amenable to Response-Guided Titration".

only have high on-target engagement and low off-target engagement. Many other factors play a role: stability, ease of manufacture, cost of manufacture, reproducibility of manufacture, ease of administration, and favorable pharmacokinetic (ADME) profile.

One important side note is that side effects may also arise from on-target engagement. As an example, a defect in the MAP/ERK pathway will lead to uncontrolled cell growth and kinase inhibitors (e.g. sorafenib)[2] inhibit this pathway to combat cancer. However, the MAPK pathway mechanism for cell growth is present throughout the body and inhibiting this pathway will typically lead to side effects in organs with rapid cellular turnover, resulting in hair loss, skin rash, and bleeding gums. Therefore, an additional challenge in drug design may be to deliver the drug to the site of action, and limit exposure in other areas of the body.

Once a viable drug molecule, administration route and dosage formulation have been identified (see also Figure **1**), the drug needs to be brought to clinical use. As a first step, preclinical research uses in vitro and animal testing to investigate safety and efficacy. Cultured cell lines are exposed to the drug in vitro, quantifying a spectrum of the toxicity endpoints. These range from simple cytotoxic effects to functional endpoints and -omics-based readouts. Proof of drug mechanism may also be shown, although a simple system may not be relevant to show effect in humans. Animal models are used because they more closely relate to humans, allowing more complex toxicity or efficacy interactions to be studied.

[2] Wilhelm et al., "Discovery and Development of Sorafenib".

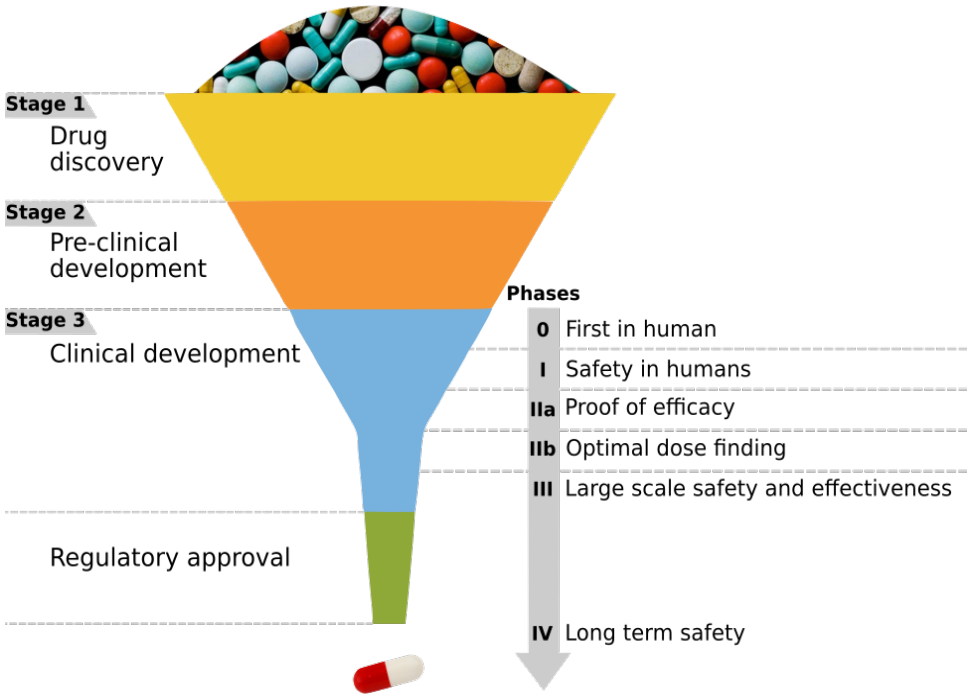


***Figure 1**: Overview illustration of the stages in drug discovery and development.*

The drug is then brought to humans. Classically, clinical research starts with Phase I, in which healthy volunteers are exposed to the drug to investigate what the maximum permissible exposure can be. These first-in-human trials start at very low doses, to maximize safety in case of unforeseen side-effects.[3] In phase II, patients are exposed to the drug, proving the drug has efficacy at a relatively safe dose level. In some cases, multiple phase II trials are used to refine dosing for patients. Finally, in phase III, the drug is tested in a large patient population, identifying rare side effects and further refining dosing. When sufficient evidence is collected to show the drug, when used as described by the label, is safe and effective, regulators may approve access to the market for the candidate drug. Post-market research may be performed to further optimize dosing regimen, or to investigate using the drug in special

[3] Bégaud, "The Rennes Disaster".

populations, such as pediatric populations. A plan to investigate appropriate pediatric dosing is often a requirement to gain market access.

A recurrent question throughout clinical research is the dose. How can we find an answer in the most cost-effective way to answer this question, without simply testing several dosing schedules?

1.2. Modeling

Modeling provides crucial information to aid in dose selection. Conceptually, we model longitudinal data using two distinct processes: pharmacokinetics (PK, what the body does to the drug) and pharmacodynamics (PD, what the drug does to the body). In PK, we describe how the drug is absorbed, distributed, metabolized and excreted. In PD, we describe how drug concentrations affect biological processes within the body. PK and PD are modeled in a similar way, assuming a system of well-mixed compartments and transport between these compartments. Pharmacometrics models are composed of three components: structural model, statistical model, and covariates model.

Two complementary approaches exist to develop the structural model: physiologically-based modeling and empiric modeling. Physiologically-based modeling (PBPK, see Figure **2**) uses fundamental research to model the body bottom-up: the body is a system of many different compartments, each corresponding to a specific organ or site in the body. Blood flows between these compartments, allowing the exchange of drug. At specific organs, the drug may be metabolized or excreted. This system of ordinary differential equations requires many fundamental parameters: typical blood flow to each organ, exchange rates, binding affinity, etc. These parameters are identified through system-specific knowledge of the body, physicochemical properties of the drug, and other (typically) in vitro or animal experiments. They are difficult to identify in humans, as individual organ concentrations are typically difficult to measure. Each parameter has meaning, allowing these models to predict other circumstances than what was originally studied. A key application is translational modeling from animal to human. These models are used to provide ball-park predictions and are a good way of predicting drug-drug interactions, first-in-human doses, or minimally effective exposures.

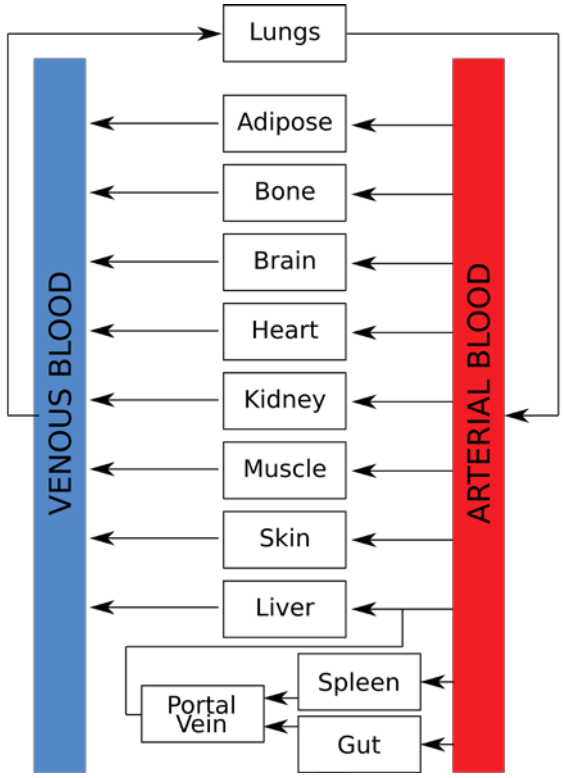


***Figure 2**: In PBPK models, every physiologically relevant organ is a separate compartment with associated size (volume) and blood perfusion rate.*

This is contrasted with empiric modeling (illustrated in Figure **3**), where human experimental data in the form of concentration-time profiles is fitted to a parsimonious model. This statistical principle of parsimony, dictating that the model should be as simple as possible, results in good descriptions of the data with models consisting of a limited number of parameters. These models are often a better fit for the data collected during clinical development, but typically only allow interpolation. Some in-between approaches exist, where semi-mechanistic modeling allows parsimonious models with parameters that still relate to physical reality.

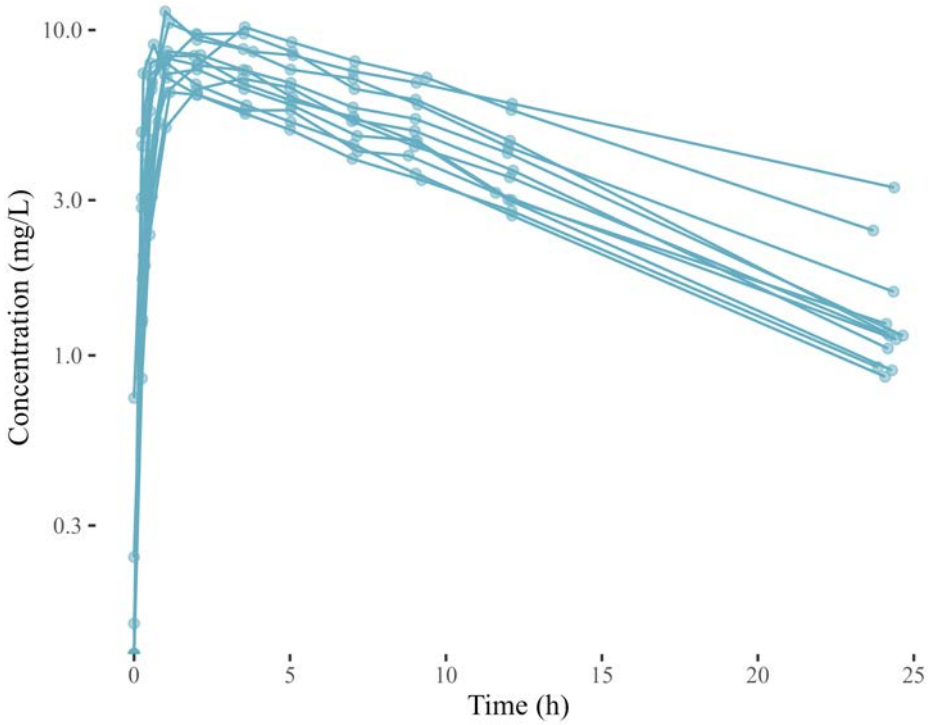


***Figure 3**: Concentration-time curves for 12 patients receiving theophylline (example dataset). A fast oral absorption and linear clearance is apparent, with no discernible distribution phase. An empiric model would therefore consist of a central compartment and an oral absorption process.*

No model is perfect. The difference between model prediction and the observed values is called *residual error*. Some residual variability may be an uncircumventible reality in the form of e.g. assay error, dosing error, and other variability that we cannot predict. However, residual error may also be the result of model misspecification. Residual error should generally be as low as possible, ideally equal to the (a priori known) measurement error. A key insight to improve pharmacometric models is to recognize there is structure in the data being modeled: observations are grouped together based on the individual patient they relate to. By introducing this *mixed effect* structure in the form of a statistical model, we can allow different parameters to be assigned to different individuals. This matches reality: different people absorb, distribute, and eliminate drugs differently, leading to different PK for the same dose. People will also respond differently to similar drug exposures,

and where an exposure may be safe in some individuals, it may be toxic in others. Nobody is average, we all differ.

Several numerical algorithms exist to estimate these parameter's typical value and inter-individual variability, optimizing the fit to a given clinical dataset. These algorithms all focus on finding the typical parameter Θ, inter-individual variability Ω and residual error Σ with optimal overall likelihood to predict the observed data. The model fit is evaluated using *Goodness of Fit* techniques. This theoretical basis and its implications on model-informed precision dosing is explored in detail in 0 of this thesis.

Predictive power can be further optimized by identifying covariate effects. As an example, body size will typically bias drug clearance, with larger patients having a higher clearance. Adapting the typical clearance based on bodyweight will explain some of the previously unexplained between-subject variability and may refine and improve the fit for non-normally distributed covariates.

Whether a model fits the data well can be evaluated through several diagnostic tools. The current gold standard is a Visual Predictive Check. In this exercise, parameters are resampled using monte carlo sampling several times for every individual subject. The resulting simulated dataset of (typically) 1000 samples per subject shows an overview of all possible outcomes for the presented clinical trial. The model is deemed a *good fit* if the confidence interval for simulated median, upper and lower quantiles matches the observed median and outer quantiles. A well-fitting model already provides interesting information: identified PK parameters such as clearance may inform steady-state dosing, and the model can quantify between-compound differences.

Pharmacometric models become truly useful once they are used to answer new questions during clinical development. Through simulation, we can explore new scenarios and inform decision making. Deterministic simulations can show how a typical patient would respond to a given dosing regimen. In stochastic simulations, we generate a large virtual population of patients using monte carlo sampling. We may administer a candidate dosing regimen to these patients and explore exposure and response distributions. This is used regularly in industry to identify candidate dosing regimen to be tested in prospective trials. During regulatory approval, it may also be used to explore whether dosing should be adapted based on relevant covariates. Population simulation may also be used to inform an initial dose for use in the pediatric population. Finally, it can be used to explore what-if scenarios,

e.g. at what level of anti-drug antibody occurrence would a candidate biosimilar lose bio-equivalence to the originator.

Further refinement is possible through clinical trial simulation. In this case, we sample only a limited number of patients and simulate realistic constraints: only a limited number of observations may be taken, patients drop out, and statistical analysis on simulated trial data is performed as if it were a real trial. This exercise is repeated many times to find the statistical distribution of clinical trial endpoints. This may be used to optimize clinical trial design. This further improves efficiency during clinical development. Over the past 20 years, these principles have been widely embraced by industry. FDA has released guidelines for Model-Informed Drug Discovery and Development (MID3) and actively encourages the use of modeling at every step of development. Pharmacometrics has improved efficiency of clinical trials and in some cases has even made clinical trials superfluous.

1.3. Model-informed drug development

Originally, dose finding closely followed clinical drug development. Phase I testing explored the maximum tolerable dose, Phase II aimed to demonstrate a drug effect at this maximum dose, and Phase III refined subgroups as required. This research strategy is appropriate for drugs where toxic exposure occurs at much higher levels than required for efficacy. We describe this as the *therapeutic window*. When a drug at dose X induces toxic side-effects in some patients, while other patients show lack of efficacy, dosing this drug is challenging. For a large number of these drug candidates, drug development was terminated early. Some compounds may instead be dose-individualized by titration. Many antipsychotic drugs are uptitrated for effect. Anticoagulants such as warfarin are uptitrated based on blood clotting tests. For many drugs though, dosing individualization is neither practical nor feasible. These drugs, while potentially promising, never made it past Phase II.

Fortunately, model-informed drug development (MIDD) has improved decision making during clinical development. Translational modeling has enabled description of compounds from pre-clinical data, informing a safe FIH dose. The phase I data can then be used to inform a PK model, and potentially even a PD model describing effect on disease-related biomarkers. This is used to inform the phase II dosing regimen, already optimizing for maximal clinical utility, with separate dosing candidates if required. Confirmatory phase III trials allow the inclusion of covariates into these models. This allows us to explore whether dose adaptation in subgroups is

required for optimal exposure. MIDD has enabled industry to find practical dosing regimen for population subgroups when required.

But what about the individual patient? MIDD does not match the original vision of pharmacometrics: optimize dosing for individual patients. Non-linear mixed effects models were originally introduced by Sheiner and Beal[4] in the late 70's. By using an iterative algorithm, they managed to identify both typical parameter values and the associated between-subject variability in the population. This was used to estimate individual parameters for new patients using only limited concentration samples, allowing accurate predictions of the future concentration-time profile and thereby identifying the optimal safe and effective individual patient dose. Is pharmacometrics nothing more than a tool to identify fixed dosing, with outliers fallen by the wayside?

Holford and Buclin[5] proposed a quantitative method for analyzing the necessity of personalized dosing. In this approach, PK exposure and associated PD outcomes for a given candidate dose are simulated. If a large portion of the PD outcome distribution is safe and effective, a fixed dosing is appropriate. If the variability is too large, it may be reduced by adapting the dose based on covariates (moving between-population variability from the endpoint into the different individual doses). Further adaptation is possible by administering a dose and measuring outcomes. These allow to identify the individual patient PK/PD parameters, allowing dose adaptation. A simple alternative may be to uptitrate-to-effect, or downtitrate-for-safety. This is regularly done, especially for over-the-counter drugs.

While useful, the drivers for precision dosing are more complex than only this theoretical approach. There are both scientific, logistic, and practical challenges for precision dosing. The following chapter will further refine these challenges.

[4] Sheiner et al., "Forecasting Individual Pharmacokinetics".

[5] "Safe and Effective Variability—A Criterion for Dose Individualization".

1.4. Precision dosing: we are not there yet

1.4.1. The hype of precision medicine: what is precision dosing?

Precision medicine or personalized medicine is an umbrella term which covers many topics. In this short section, we provide an overview of these topics, and position model-informed precision dosing in this space.

A first domain is the development and production/synthesis of compounds specifically targeted to individual patients or groups of patients. There is a gradient from *can work for anyone* to *specifically designed for only patient X*. It is obvious that compounds targeting common disease pathways -effecting millions of humans- can hardly be called "precision medicine". In contrast, a chemotherapy drug like erdafitinib[6] may only be administered to combat FGFR-mutated cancer cells, requiring a biopsy for confirmation of susceptibility. Going further, some drugs are *grown* to specifically target the DNA of a single patient. Such is the case for recombinant adeno-associated virus gene therapy (e.g. Zolgensma, Luxturna, Glybera).[7] The development of these types of 'personalized medicine' or 'targeted therapy' is not what we will focus on in this work.

Another approach within this umbrella term is 'personalized healthcare,'[8] the use of smart apps and devices to track patient behavior in realtime and suggest healthcare options. A prime example of its use during drug development is the tracking of Parkinson's Disease severity using a smartwatch[9]. This field is still in its infancy, and the connection to population pharmacokinetic/pharmacodynamic modeling remains unclear. At the same

[6] Dosne et al., "Erdafitinib's Effect on Serum Phosphate Justifies Its Pharmacodynamically Guided Dosing in Patients with Cancer".

[7] Keeler and Flotte, "Recombinant Adeno-Associated Virus Gene Therapy in Light of Luxturna (and Zolgensma and Glybera)".

[8] Andreu-Perez et al., "From Wearable Sensors to Smart Implants--Toward Pervasive and Personalized Healthcare".

[9] Powers et al., "Smartwatch Inertial Sensors Continuously Monitor Real-World Motor Fluctuations in Parkinson's Disease".

time however, smartphones, smartwatches and other wearable technology (e.g. insulin pumps!) are an ideal platform to deploy model-informed precision dosing software in close proximity to patients and physicians.

Looking at statistical techniques in this space, the term *personalized medicine* is also used to describe a method identifying the optimal diagnosis/disease for an individual. For many diseases, plenty of treatment options are available. While selecting the optimal treatment for an individual patient was historically a task for physicians, statistical techniques use causal inference[10] to predict treatment benefit and risk based on covariates available at the time of decision.

> "Personalized medicine is the ability to use an individual's genetic make-up and life experiences to diagnose and treat disease. […] If the response to treatment is a unique characteristic of the individual that cannot be predicted a priori, then true personalized medicine has little practical utility in medicine or biomedical research." - Yazdani et al, 2015

Let us focus precisely on the situation where *the response for a candidate treatment cannot be predicted a priori*. Linking back to Holford and Buclin[11], we know MIPD focuses on cases where unexplained inter-individual variability is so large that dose adaptation based on ongoing response is needed. MIPD employs the up-to-now a posteriori exposure and/or response to a treatment to optimize this ongoing treatment.

Both personalized medicine and MIPD are complementary. MIPD can be considered as an optimization technique to increase benefit-risk ratio, akin to the switch from skin patch to chewing gum for medicinal nicotine delivery. The overall benefit-risk ratio for a treatment using MIPD may be described either through a clinical trial, or through simulation of MIPD for a given patient. As such, MIPD provides causal inference techniques for personalized medicine with a (more) stable response to treatment that can be better predicted a priori.

[10] Yazdani and Boerwinkle, "Causal Inference in the Age of Decision Medicine".

[11] "Safe and Effective Variability—A Criterion for Dose Individualization".

1.4.2. Scientific challenges

Precision dosing faces several scientific challenges, which are rooted in the transition from population prediction to individual prediction. Population modeling is well studied, and we know how to build and validate models that fit the data, and that make realistic predictions. This fundamental methodological research is missing for MIPD.

A first area of active research focuses on how population models can be used for precision dosing. The research group at UHamburg[12] explored the prediction accuracy and precision of existing population models. They found large differences between model predictions, showing that evaluating a model is required before an existing model is used for precision dosing. Hughes and Keizer[13] at InsightRx have explored downweighing the importance of population priors, as these may not be appropriate for the current patient. Flattened Priors increase the measured inter-individual variability and thereby allow individual parameters that are more divergent from the original population, in favor of a closer model fit to the observed patient data.

The idea of flattened priors is related to an idea from 10 years ago: non-parameteric modeling. In the introductory paper from Neely et al,[14] the group argue that forcing inter-individual variability to fit a particular distribution will always entail compromise and error. Instead, non-parametric modeling uses an adaptive grid to characterize the likelihood of individual patient parameters. While the basic premise is different, both Keizer and Neely argued the same thing: log-normal inter-individual variability as estimated by NLME should not be taken as scripture, but rather be flexibly interpreted.

---

[12] Broeker, "Towards Precision Dosing of Vancomycin".

[13] "A Hybrid Machine Learning/Pharmacokinetic Approach Outperforms Maximum a Posteriori Bayesian Estimation by Selectively Flattening Model Priors".

[14] Neely et al., "Accurate Detection of Outliers and Subpopulations With Pmetrics, a Nonparametric and Parametric Pharmacometric Modeling and Simulation Package for R".

A second area focuses on unexplained parameter drift. We define parameter drift as any occurrence where a parameter changes at a specific timepoint in an unexplained but non-random way. Clinical examples are plentiful: the patient recovers from surgery and is no longer bedridden, the patient receives co-medication with a drug-drug interaction, the patient gets an unrelated severe infection resulting in increased cardiac output, etc. Some of these occurrences are investigated during drug development. Multi-arm PK studies comparing fasted and fed administration of oral drugs are commonly performed during Phase I. Common co-medication causing DDIs are also investigated, although this is often limited to a qualitative description. In some cases though, these effects can be incorporated in the model. To optimize dosing of meropenem in a 19-month old female, Saito et al.[15] used a pharmacokinetics model where clearance depended on renal dialysis flow rate.

Lacking an explanatory covariate though, these transient variations are either described by residual error, or by parameters varying randomly between occasions (inter-occasion variability). Both types of variability are random, while the above description is anything but random. Abrantes et al.[16] recommended to determine the next dose by discarding IOV estimates, however this was based on simulated data, not clinical data with non-memoryless variability. Karlsson, Beal, and Sheiner[17] showed a residual error model with autocorrelation may be more appropriate to model this data. However, this approach simplifies parameter drift into residual error. A pragmatic solution is offered by Keizer et al.[18] by adding increased weight to more recent observations.

Theoretically, stochastic differential equations (SDE) may be the most appropriate method to model unexplained parameter drift. This allows parameters to drift randomly over time, implemented as Brownian motion.

---

[15] "Meropenem Pharmacokinetics During Extracorporeal Membrane Oxygenation and Continuous Haemodialysis".

[16] "Handling Inter-occasion Variability in Model-based Therapeutic Drug Monitoring".

[17] "Three New Residual Error Models for Population PK/PD Analyses".

[18] "Model-Informed Precision Dosing at the Bedside".

Leander et al.[19] demonstrated SDE allows to distinguish between residual error and model dynamics. Estimation of SDE models has only recently been implemented in NLME estimation software, and modeling methodology has not yet been firmly established. To the best of our knowledge, SDE has not been applied to precision dosing yet.

1.4.3. Industry challenges

This section was based in part on the work by Peck[20] and Lesko[21]

Over the last 10 years, a steady movement to model-based "Learn and confirm" has taken place in clinical drug development. Whereas regulators were satisfied with a Phase 3 trial showing clinical superiority over either placebo or standard of care in the past, they are now demanding proof that the drug is administered in the most optimal way for all patients[22]. In an ISoP position statement[23] from 2018, exposure-response modeling and individualized dosing were considered key components of any submission dossier.

Why then were 86%[24] of FDA-approved drugs between 2013 and 2017 marketed on a fixed dose schedule? Logic dictates this is either because drugs requiring dose individualization are rare, or because these drugs never make it to market. Wang et al found that only 64% (76 of 119) of approved drugs were amenable to response-guided titration, although potentially more drugs are amenable to therapeutic drug monitoring. Of these 76, FDA

---

19 "Mixed Effects Modeling Using Stochastic Differential Equations".

20 "Precision Dosing".

21 "Perspective on Model-informed Drug Development".

22 Zhang et al., "Exposure–Response Assessment in Pediatric Drug Development Studies Submitted to the US Food and Drug Administration".

23 Maloney et al., "Comment from International Society of Pharmacometrics on Exposure-Response Analysis in Drug Development and Regulatory Decision Making; Request for Comments (Docket No. FDA-2018-N-0791)".

24 Wang et al., "A Systematic Assessment of US Food and Drug Administration Dosing Recommendations For Drug Development Programs Amenable to Response-Guided Titration".

recommended efficacy trials evaluating individualized dosing for 35 compounds, with only 11 compounds finally having individualized dosing regimen in the label. Wang et al. argue the systematic review shows FDA-sponsor interactions are more likely to enable individualized dosing strategies in labeling, yet only a mere 11 of 76 drugs finally included dose individualization instructions in the label.

Dosing individualization is rarely developed by industry, especially when fixed dosing may also allow market access. This is not surprising, as industry challenges are plenty. For a new compound requiring dose individualization, a validated accompanying assay is needed. Any computer program for dosing individualization is considered as a medical device, requiring European CE approval or FDA Software as a Medical Device 510(k) notification. Market access and logistical distribution of the drug becomes much more complex.

Would industry not benefit from dose individualization leading to a far superior treatment? Unfortunately, this superiority is far from certain, and this uncertainty destroys all incentive to overcome the aforementioned roadblocks. There has been little research into predicting the performance or outcomes of precision dosing. Buclin et al.[25] outlined how to decide on the usefulness of precision dosing, but failed to capture the great investment risk of showing value in clinical trials. A company would need solid evidence of the value and market potential of a compound requiring precision dosing.

Such efforts are rare however, and most clinical trials into precision dosing are sponsored by academic centers on approved drugs. A search of public trial registries was performed (full results in Appendix Appendix C). We searched for trials including 'precision dosing' or 'computer dosing' in the intervention, yielding 42 trials. Only clinical trials where dosing was adapted using a computer algorithm were retained, this yielded 33 trials. None were industry-sponsored. All associated publications were reviewed for a description of a priori statistical power. Of the 7 studies with published protocols, two admitted to no formal power calculation at all. The other studies either applied a guesstimate, or used endpoints from historic control. In a single trial (without published protocol), Hughes et al.[26] characterized

[25] "The Steps to Therapeutic Drug Monitoring".

[26] "Bayesian Clinical Decision Support-Guided Versus Clinician-Guided Vancomycin Dosing in Attainment of Targeted Pharmacokinetic Parameters in a Paediatric Population".

outcomes for vancomycin MIPD using a placebo historic control and simulated MIPD arm. This unique exercise was performed in collaboration with InsightRx, a developer of MIPD software.

In summary, we found no industry-sponsored clinical trials investigating MIPD. Academic trials investigate MIPD using enrollment sizes based on guesstimates. Quantitative power calculation is feasible but remains under-used, and as such does not reduce uncertainty in industry development of compounds requiring MIPD, in light of the other challenges MIPD poses.

Regulators attempt to tip the scales. Whereas demonstrating a population benefit seemed sufficient in the past, requirements for evidence of optimal dosing in each individual patient are rising. Healthcare payers have suggested value-based pricing to force industry in developing treatments that are used more optimally. Recent years have also seen updates to software medical device regulations, making the pathway clearer at least -but often not more accessible per se.

1.4.4. Clinical challenges

Finally, precision dosing is not easy to bring into clinical practice. Even if we assume a validated assay, a pharmacometric model, and appropriate dosing optimization software are available, then administering a fixed dose is still preferable in terms of physician workload.

First, a large amount of patient data is required as compared to fixed or bodyweight-based dosing. An accurate dosing history, observation history and covariate overview are required by the software. Either they are entered manually -a daunting task for drugs administered daily-, or they are collected from an EHR. The latter case requires software development, as even standardized systems for data transfer are often not suited for the transfer of pharmacometric data. These EHR data require transformation to be suited as input for pharmacometric models, and this requires business knowledge.

Software deployment is tricky. These are safety-critical systems and should work reliably. As this is patient data, the General Data Protection Regulations (GDPR, EU regulations) and Health Insurance Portability and Accountability Act (HIPAA, US regulations) apply. It is unfeasible to fully anonymize pharmacometric data. The data should therefore not leave the hospital zone of control. Computer servers should be hosted internally and be well protected.

Third, any precision dosing approach should work seamlessly within the established clinical workflow. Precision dosing should only increase workload if a clear associated benefit is present. Each hospital service will have different workflows, and careful change management and software adaptation is needed. Physicians may not be convinced of the value, and indeed evidence of true clinical value is often missing for precision dosing approaches, showing only improvement in PK exposures as clinical trial enrollment numbers to show clinical improvement is unfeasible.

Chapter 2 Nonlinear mixed effects modeling and precision dosing

This chapter details the theoretical foundation used to perform MIPD. It serves as a mathematical, numerical and practical introduction, summarizing all background information in a single book chapter. Readers experienced with NLME and MIPD are invited to gloss over this chapter.

For a more in-depth view, we refer to the work of Hedaya on clinical pharmacokinetics,[27] and the comprehensive introduction to NONMEM modeling by Owen and Kelly.[28] For more advanced users, the excellent tutorials[29] by J. Bauer provide a comprehensive starting point for newer estimation techniques such as MCMC, SAEM and mu-referencing.

2.1. Non-linear structural models

Mathematical equations Models essentially translate basic assumptions about biology into mathematical equations. In the case of pharmacokinetics or pharmacodynamics models, we assume a system of interconnected well-mixed compartments. Compartments are represented by an amount, usually of a drug or other biologically relevant biomarker. PBPK models have compartments that match actual organs, and transfer rates between compartments are calculated from experimental observations or mechanistic models. Empirical models simplify these systems into the minimal number of parameters required, dependent on the shape of the curve the model is fitting. Compartments, parameters and transfer rates do not necessarily match biological processes one-to-one.

---

[27] Hedaya, *Basic Pharmacokinetics*.

[28] Owen and Fiedler-Kelly, *Introduction to Population Pharmacokinetic/Pharmacodynamic Analysis with Nonlinear Mixed Effects Models*.

[29] Bauer, "NONMEM Tutorial Part I"; Bauer, "NONMEM Tutorial Part II".

An example of a simple 1-compartment system with intravenous (IV) administration and linear clearance is given below:

$$\frac{dA1}{dt} = -Ke \times A1$$

$$A1(t = 0) = D$$

$$CONC = \frac{A1}{V1}$$

The amount of drug $A1$ in the central compartment 1 is set to the dose D at time 0. It then evolves over time, diminishing by $Ke * A1$ per hour. Ke is the elimination rate constant. As biological systems frequently describe elimination as a volume flow, this may also be restated using the clearance CL as $-Ke \times A1 = -CL \times \frac{A1}{V1}$.

Numerical solving Systems of differential equations can be solved numerically, and many different methods exist. Below, we give the basic mathematical principle behind these methods and why they work. The methods most frequently used are Runge-Kutta (RK) solvers, Adams-Bashforth (AB) solvers, and backward differentiation formula (BDF) solvers. In RK solvers, the ordinary differential equation system is approximated by a straight line to calculate all values for a time $t_{i+1} = t_i + h$, with h the chosen step-size. When the differential equation is given as $\frac{dy}{dt} = f(t, y)$, 1st order RK (denoted as *RK1*) is as simple as

$$y_{i+1} = y_i + h * f(t_i, y_i)$$

In other words, the function is approximated by a line, with the slope given by the (known) derivative function. We should pause for a brief moment to introduce some key concepts of numerical methods. The work required by the above method is small, as only one evaluation of $f(\quad)$ is required. The work required to calculate $y(t)$ is directly related to the stepsize. A higher stepsize allows to advance faster, yet increases numerical error dramatically. To reduce error, one could either take smaller (but more) steps, or perform more accurate steps. Accuracy can be improved by doing more work each step.

2nd-order RK uses RK1 to calculate a temporary value k_1, which is the value of the function at half the step-size h. It then uses the derivative at the point as derivative for the full step h.

$$k_1 = y_i + \frac{h}{2} * f(t_i, y_i)$$

$$y_{i+1} = y_i + h * f\left(t_i + \frac{h}{2}, k_1\right)$$

Higher-order RK methods iterate on this principle further by adding additional intermediate points. RK methods are usually implemented as tandem p and $p - 1$-order methods. This allows an estimate of the integration error. We can adapt the step-size automatically to ensure this error remains below a given maximum value. Adams-Bashforth methods are similar to RK methods, but use a polynomial to approximate the target function over the past N points. This results in a higher computational cost per step, gaining increased accuracy per step.

Some ODE systems are highly sensitive to small changes in initial value. We call these *stiff* systems. If a small error is introduced, this error is compounded and can lead to highly different outcomes. Although RK and AB methods are fast, they generally perform badly for stiff systems. In this case, BDF methods can be used. Instead of directly calculating the next value y_{i+1}, they solve the equation $p'(t_{i+1}) = f(t_{i+1}, y_{i+1})$, with $p(\quad)$ the polynomial fit through the last N calculated points. BDF methods are computationally more expensive, but much more stable. To solve both stiff and non-stiff systems, Hindmarsh and Petzold[30] developed the LSODA solver. This solver switches between stiff and non-stiff methods automatically, and is therefore an ideal candidate for solving user-specified ODE systems without requiring numerical insight into stiffness.

Numerical solving of ODE systems works well, but requires a lot of calculation. Fortunately, closed-form algebraic solutions can be derived for simple ODE systems. They are implemented in most pharmacometrics software, resulting in a massive computational speed boost.

Application to pharmacometrics Pharmacometrics models are written as algebraic equations, or as systems of differential equations. The choice of an appropriate solver and numerical precision is important to keep calculation

[30] "LSODA, Ordinary Differential Equation Solver for Stiff or Non-Stiff System".

times low. An example of simulation of individual profiles is provided in Figure **4**

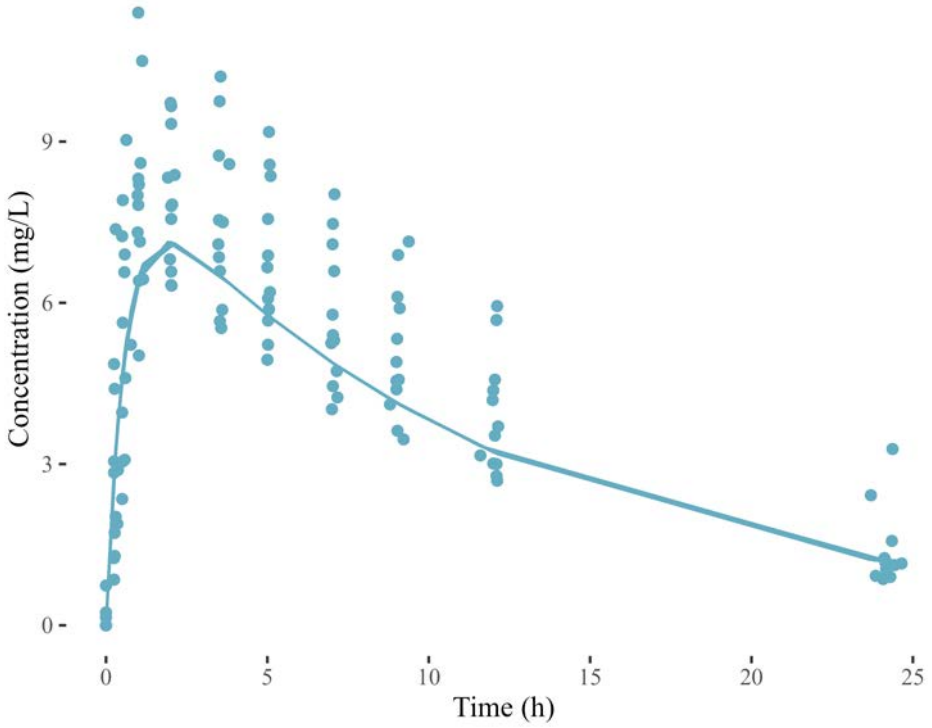


Figure 4*: Concentration-time profile prediction using non-linear model with parameters KA=1.7, V=0.6 and CL=0.05.*

2.2. Fitting non-linear models

Mathematical equations The solution to the ODE system can be written as $f(t_i, X) = Y_i$, with X the individual pharmacometric parameters and Y_i the predicted values in a patient at times t_i. To fit this model to patient measurements DV (Dependent Variable), we introduce a residual error term ϵ.

$$DV_i = Y_i + \epsilon_i$$

$$\epsilon_i \sim \mathcal{N}(0, \sigma)$$

ϵ is centered around 0 with standard deviation σ. We must now find the most likely set of parameter values X, given the observed values DV_i. This likelihood $p(X|DV)$ cannot be directly calculated, but we can apply Bayes rule to transform this.

$$p(X|DV) = \frac{p(DV|X)p(X)}{p(DV)} \qquad (1)$$

As the demoninator $p(DV)$ does not depend on X, it can be dropped. There is no prior distribution for $p(X)$, so this can be dropped as well.

$$p(X|DV) \sim p(DV|X)$$

The likelihood of all n datapoints DV being predicted by parameter values X is then given by

$$p(DV|X) = \prod_i^n p\,(DV_i|X) = \prod_i^n \Phi\,(DV_i - Y_i, \sigma)$$

We can find the maximum probability using numerical optimization. In practice, the 2-log-likelihood is used. This is easier to calculate, and the maximum of both functions occurs at the same parameter value X.

$$2 * log\big(p(X|Y)\big) = 2 * log\left(\prod_i^n \Phi\,(DV_i - Y_i, \sigma)\right)$$

$$= 2 * \sum_i^n l\,og\big(\Phi(DV_i - Y_i, \sigma)\big)$$

$$= 2 * \sum_i^n l\,og\left(\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{DV_i - Y_i}{\sigma}\right)^2}\right)$$

$$= -\sum_i^n \left(\frac{DV_i - Y_i}{\sigma}\right)^2 - 2 * \sum_i^n l\,og\big(\sigma\sqrt{2\pi}\big)$$

$$= -\sum_i^n \left(\frac{DV_i - Y_i}{\sigma}\right)^2 - 2n \cdot log\big(\sigma\sqrt{2\pi}\big)$$

As the last term is not dependent on the datapoints, it can be omitted.[31] Minimizing this function -a straightforward weighted least squares problem- allows us to find parameter values that fit $f(t_i, X) = Y_i$ as close to DV_i as possible.

Numerical solutions Finding the optimal set of parameter values is performed using numerical optimization algorithms. In this section, we will discuss two general approaches: gradient/hessian-based approaches, and the Nelder-Mead method[32]. Finding the minimum of a function is intuitively parallel to finding the lowest point on a surface.

Generally speaking, the minimum of a function $f(X)$ is the set of parameter values $\hat{X}$ for which the first derivative $f'(\ \ )$ (also called the *gradient*) is 0, and the second derivative $f''(\ \ )$ (also called the *hessian*) is positive. Gradient-based methods calculate the gradient and move the estimate for $\hat{X}$ in the direction of the gradient. This is similar to a kangaroo searching the bottom of the valley by jumping in the steepest direction downward. The size of each jump depends on the curvature, with long hops more appropriate for flatter surfaces. A relatively modern[33] version is the Broyden–Fletcher–Goldfarb–Shannon (BFGS) algorithm[34]. The algorithm uses only gradient calculations, but gradually approximates the hessian at each step. This algorithm works best when an equation for $f'(\ \ )$ is available. The gradient can be approximated by $f'(X) = \frac{f(X+\epsilon/2)-f(X-\epsilon/2)}{\epsilon}$, this is known as the finite differences method. Such a method requires high precision on $f(\ \ )$, so this is costly to evaluate, especially when $f(\ \ )$ is the result of ODE integration.

Alternatively, we can rely on only function evaluations. The Nelder-Mead method[35] forms an n+1-simplex when searching in n-dimensional space (e.g. a triangle when searching in 2-dimensional space). Each step, we

[31] Pinheiro and Bates, "Approximations to the Log-Likelihood Function in the Nonlinear Mixed-Effects Model".

[32] Bultheel, *Inleiding Tot de Numerieke Wiskunde*.

[33] Sir Isaac Newton developed a method for numerical optimization in the 17th century.

[34] Zhu et al., "Algorithm 778".

[35] Nelder and Mead, "A Simplex Method for Function Minimization".

intuitively let the triangle tumble down the surface until it lies flat. At that point, we know the minimum is contained within the simplex. We can shrink the simplex so it can further tumble towards the lowest point. This method is cheaper per iteration, as it needs to evaluate $f(\quad)$ less. However, it converges slower to the desired minimum. This is a heuristic method; there is no guarantee that the method converges or will find the optimal value, especially in the case of irregular log-likelihood surfaces.

Application to pharmacometrics When fitting a model, we use numerical optimization techniques to find the most likely parameter values. A model expressed using algebraic equations is easier to fit, as the gradient and hessian can be directly calculated. Models using ODE systems benefit from heuristic optimization approaches such as Nelder-Mead, however these may not converge in the case of irregular log-likelihood surfaces. This is the case when e.g. an absorption lag-time is used in the model. An example of a non-linear model fit is provided in Figure **5**.

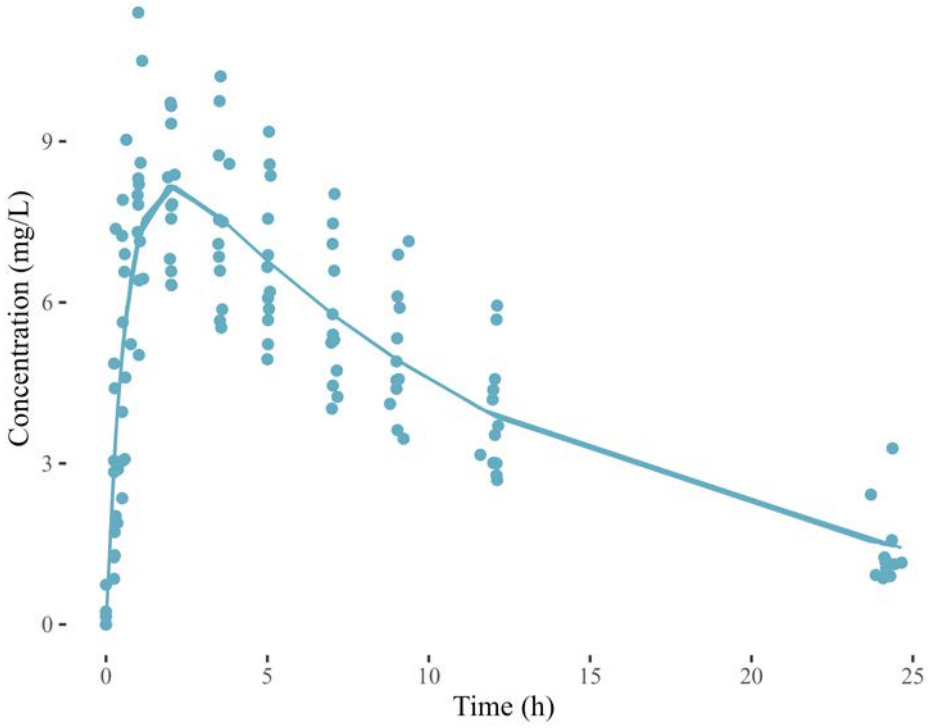


***Figure 5**: Concentration-time profile prediction using non-linear model with parameters fitted by non-linear least squares.*

2.3. Mixed effect models: prior distributions

Mathematical concept As discussed previously in 0, pharmacometric data has a unique structure: observed data can be grouped per patient. We add *inter-individual_variability*[36] (IIV) : each pharmacometric parameter is different per patient.

$$X = \theta * e^{\eta}$$

$$\eta \sim \mathcal{N}(0, \omega)$$

Mathematical equations In the equations below, we will use Θ, Ω, Σ whenever possible to represent the full matrices with correlation, instead of θ, ω, σ that are used when assuming independent distributions. Looking at equation (1) again for only patient j, we now have a prior distribution for parameters X_j:

$$\begin{aligned} p(X_j|DV_j) &= \frac{p(DV_j|X_j)p(X_j)}{p(DV_j)} \\ &\sim p(X_j)p(DV_j|X_j) \\ &\sim \Phi(\eta_j, \Omega) \times \prod_i^n \Phi\left(DV_{i,j} - Y_{i,j}, \Sigma\right) \end{aligned}$$

The above equation shows the *marginal_likelihood* for an individual η_j of patient j. Finding this estimate is called Empirical Bayesian Estimation (EBE). Intuitively, it is composed of two parts: the likelihood of encountering such a patient in the population, and the likelihood of observing these values in such a patient. The most likely parameters can be easily found using the same optimization algorithms from section 2.2. We can transform the above equation using 2-log-likelihood, assuming independent distributions for inter-individual variability. This is known as the Maximum A Posteriori (MAP) estimate.

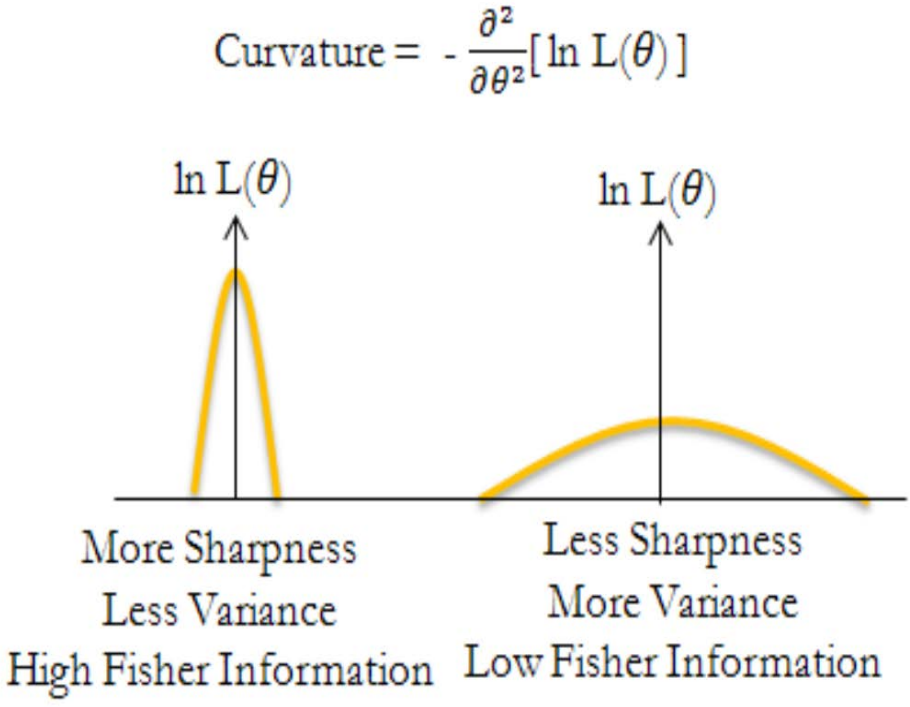
$$2LL = -\sum\left(\frac{DV_{i,j} - Y_{i,j}}{\sigma}\right)^2 - 2n * log(\sigma\sqrt{2\pi}) - \sum\left(\frac{\eta_j}{\omega}\right)^2$$

Looking at 2LL for the most likely individual parameters $\hat{\eta_j}$ (also called the *mode*), the shape of the loglikelihood function determines how certain we are

[36] Also called between-subject variability (BSV)

about these parameters. The standard error on individual parameter estimates $\widehat{\eta_j}$ is called ϕ_j. Intuitively, the more *pointy* the curve, the more precise our estimate. Formally, the observed fisher information matrix $OFIM$ determines parameter uncertainty using the following equation:

$$\phi_j^2 = var(\hat{\eta}) = OFIM^{-1} = \{-Hess\}^{-1}$$



In some cases, the hessian cannot be evaluated. The marginal likelihood can then be sampled from directly using markov chain monte carlo methods (MCMC) such as Metropolis-Hastings sampling[37].

Whether EBE estimation uncertainty is relevant, is highly dependent on the application. In many cases, clinicians are interested in the *optimal* dose to hit a certain target AUC or trough concentration. Aim for the center, and shoot.

[37] Hastings, "Monte Carlo Sampling Methods Using Markov Chains and Their Applications".

When multiple targets are applied, it is appropriate to explore the full parameter uncertainty range for a candidate dosing regimen. Consider the example in Figure **6**, where the full concentration range should be within the therapeutic window. While both dosing regimen are in target, parameter uncertainty shows a high probability of exceeding the target window for the 30mg dosing.

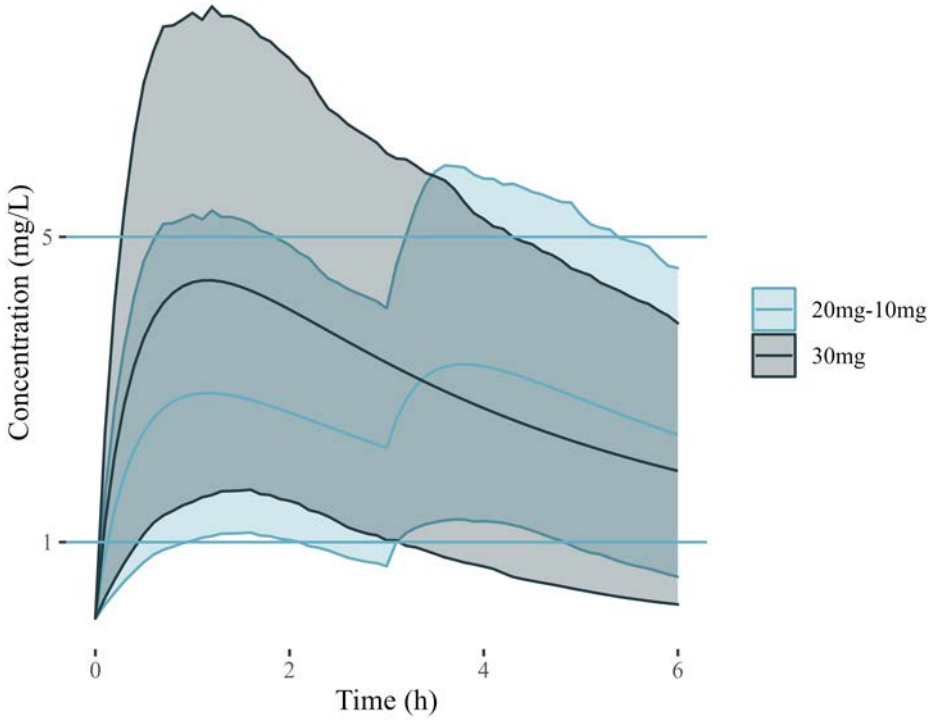


***Figure 6**: For this hypothetical drug, the therapeutic window is between 1 and 5. Without uncertainty, a single dose of 30mg would be appropriate. When considering uncertainty, the 20-10mg regimen has a higher probability of concentrations in the therapeutic window.*

2.4. Mixed effect models: estimating distributions

Equation for population The previous section assumes we know the population values θ and ω. To find the a priori likelihood for population

parameters θ and ω, we can evaluate the marginal likelihood across all possible individual η_j values for all patients j.

$$p(X_j|DV_j) \sim \Phi(\eta_j, \Omega) \times \prod_i^n \Phi(DV_i - Y_i, \Sigma)$$

$$p(\Theta, \Omega|DV) \sim \prod_j \left\{ \int_{\eta_j=-\infty}^{\infty} \Phi(\eta_j, \Omega) \times \prod_i \Phi(DV_{i,j} - Y_{i,j}, \Sigma) d\eta \right\}$$

Numerical solving As compared to the previous section, the main challenge in population model fitting is to solve the integral across all values. We will discuss two approaches: the laplacian approximation, and the SAEM algorithm.

The first approach simplifies the integral through a laplacian approximation[38] around the mode $\hat{\eta}_j$. Intuitively, we approximate the marginal likelihood by a normal distribution.

$$\int p(\eta_j|DV_{i,j}) d\eta_j \approx \int \Phi(\eta_j, \Omega) \times \Phi(\eta_j - \hat{\eta}_j, \phi_j) d\eta_j$$

As $\hat{\eta}_j$ is conditional on the estimate for Θ, Ω, an iterative process is still required, yet the individual steps are greatly simplified. Further simplifications result in the FOCEI, FOCE and FO methods. They are not further discussed here, we refer to the work by Wang[39] for more details.

A second approach is the Stochastic Approximation and Expectation Maximization algorithm[40]. In this algorithm, the marginal likelihood is not approximated. Instead, the algorithm uses metropolis-hastings sampling to draw candidate η_j from the marginal distribution for a given Θ, Ω, and uses the mean of these candidate η to update the population parameters. At each iteration, candidate η_j are replaced with a probability dependent on the marginal likelihood. This process iterates towards the optimal population values.

---

[38] Gelman, *Bayesian Data Analysis*.

[39] "Derivation of Various NONMEM Estimation Methods".

[40] Kuhn and Lavielle, "Maximum Likelihood Estimation in Nonlinear Mixed Effects Models".

Application to pharmacometrics Non-linear mixed effects models describe observed values through both residual error and inter-individual variability. Fitting NLME models is computationally expensive, so numerical algorithms are required. In Figure **7**, a single individual fit is shown. Figure **8** shows all individual fits, as well as corresponding population predictions.

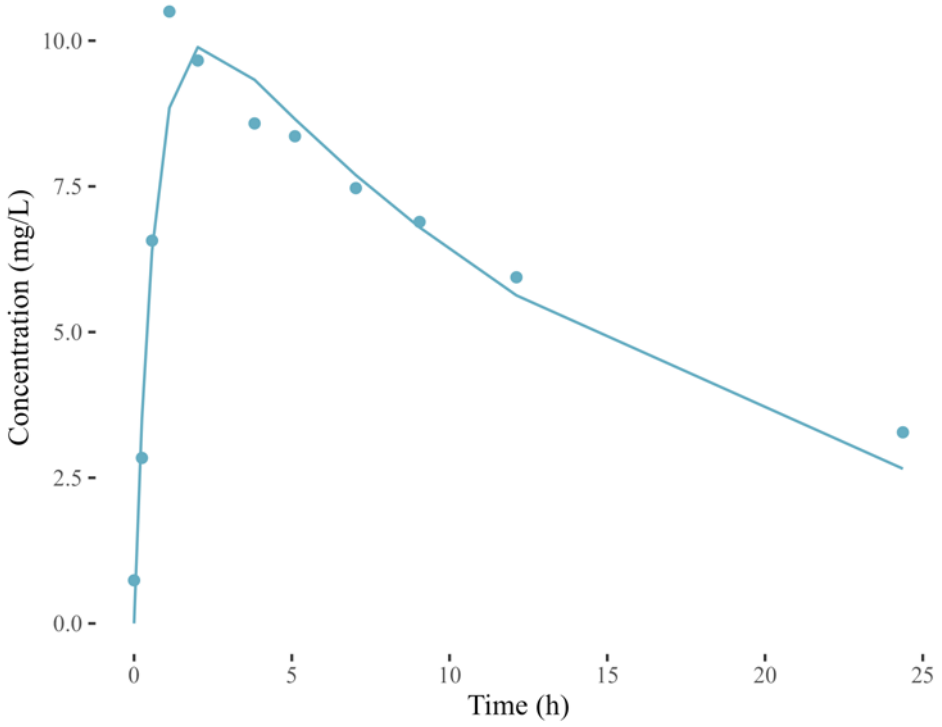


***Figure 7**: Theophylline time-concentration data for a single individual. The line shows the individual fit of a non-linear mixed effects model, with inter-individual variability estimated through the `optimr` non-linear optimization function.*

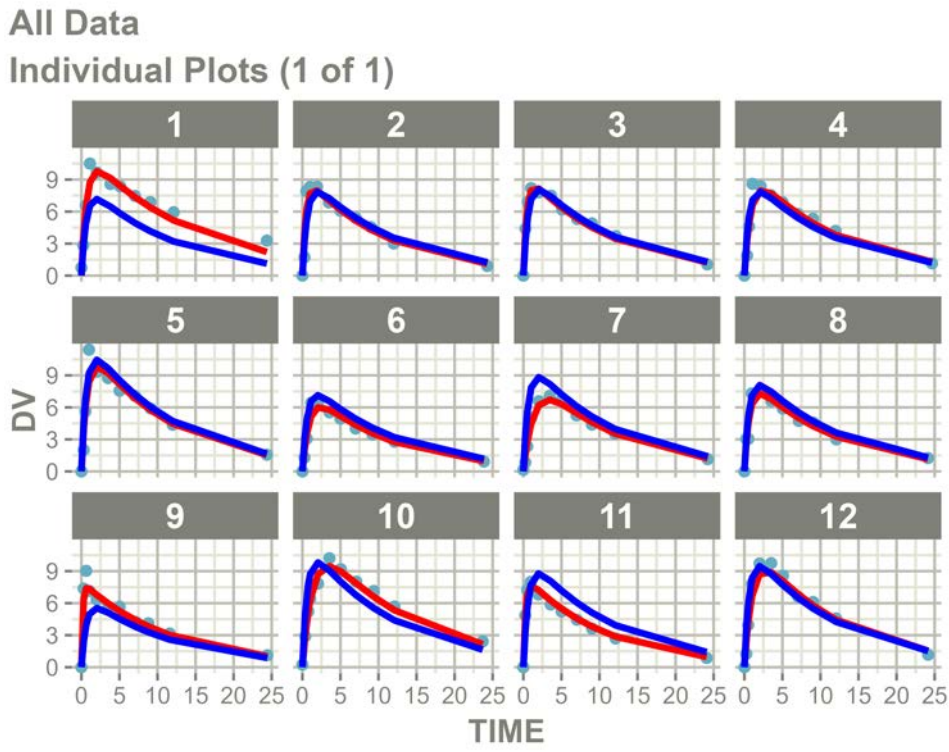


***Figure 8**: Individual concentration-time plots of non-linear mixed effects model fitted to theophylline dataset (12 patients). Population prediction (blue) and individual prediction (red), datapoints (light blue).*

2.5. Covariate models

In NLME models, unexplained inter-individual variability can be reduced by introducing covariates. These may explain certain variation to parameter values, thereby reducing unexplained IIV. In the below example, we adapt the equations for clearance and distribution volume to include bodyweight, using allometric scaling theory.[41]

[41] Holford and Anderson, "Allometric Size".

$$V = \theta_V * e^{\eta_V} * (BW/70)^1$$
$$CL = \theta_{CL} * e^{\eta_{CL}} * (BW/70)^{\frac{3}{4}}$$

We assume the covariate is well known and has a direct influence on individual parameters. An alternative method of incorporating parameters is joint modeling. We treat covariates on an equal footing with other observed values. The model predicts covariate values with inter-individual variability, and a residual error model applies to covariate observations just as with drug concentration observations.

2.6. Precision dosing

We previously established the EBE algorithm can be used to estimate individual patient parameters. We can use these parameters to predict for future doses. Precision dosing now constitutes of finding the appropriate future dose to hit a certain target. In its simplest form, this algorithm can be formulated as follows:

Step 1: Estimation

$\hat{\eta}$ = argmax $p(\eta|Y)$

= argmax $p(\eta|prior) \cdot p\big(Y|f(\eta)\big)$

Step 2: Optimization

Apply root finding on $g(D') = f\big(\hat{\eta}, t_{target}, D'\big) - C_{target}$

Step 3: Confidence interval

Sample individual η distribution using OFIM-derived variance or MCMC

Calculate $C_{pred} = f(\eta, D')$ distribution

An EBE estimation is first performed on all observed data. The estimate $\hat{\eta}$ maximizes the marginal likelihood, i.e. the likelihood of encountering this η in the population, and encountering the observed values in the model prediction $f(\eta)$. This estimate is then used to find the optimal dose D' that is

predicted to hit the target concentration C_{target} at time t_{target}. In practice, the second equation is solved by searching the root of function $f(\hat{\eta}, t_{target}, D') - C_{target}$, i.e. the dose D' for which $f(\hat{\eta}, t_{target}, D') - C_{target} = 0$. Finally, the resulting concentration using the optimal dose D' can be predicted for the full individual η distribution. This shows which concentration range can be expected with the currently recommended dose.

Three aspects may make this more complex: the search domain, the objective, and the presence of uncertainty. The *search domain* represents all possibilities for clinical intervention: dose, infusion time, formulation, dosing interval, dose suspension time. Some of these may be continuous (e.g. infused dose), while others may be discrete options (e.g. formulation, or number of 1mg capsules to administer). N continuous search options result in an N-dimensional search space, while M discrete options lead to M parallel search spaces. The *objective* can become more complex as well. In its simplest form, we target a given exposure that is correlated with efficacy and safety. However, multiple targets may exist: at least X exposure, at the lowest dose, preferring the regimen from the label, and preferring no administrations between 22:00 and 06:00. Finally, *uncertainty* may need to be incorporated. If the individual probability of target attainment is too low, different dosing modalities may need to be explored (e.g. suggest continuous infusion instead of intermittent injection to reduce risk of toxic exposures), or more frequent observations may be recommended.

While flexible precision dosing may seem enticing, results need to be effectively communicated to clinicians. The solution for a steady-state target $f_{SS}(D, Tinf, Tau) = TARGET$ in a 3-dimensional search space for dose D, infusion time $TInf$, and dosing interval Tau, including parameter uncertainty, is a fuzzy 2-dimensional surface. This is difficult to communicate to clinicians. Depending on the problem, precision dosing can therefore either take the form of a single dosing recommendation, or of a decision support tool allowing free exploration of possible solutions.

Chapter 3 Objectives

3.1. Simulate MIPD: general framework

Objective 1

Develop software and methodology to quantify model predictive performance, simulate outcomes applying MIPD to a virtual patient population, and simulate clinical trials including MIPD

Currently, mathematical software in pharmacometrics falls in two major categories: modeling tools and simulation tools. Modeling tools allow the estimation of statistical distributions of parameters, given a clinical dataset. Simulation tools use monte carlo sampling on these parameter distributions to construct a virtual population and simulate outcomes for new treatment regimen. Simulating precision dosing requires both individual parameter estimation, as well as simulation of future new treatment regimen. No flexible software currently exists to do these tasks programmatically.

We aim to explore how such a tool may be created, and postulate this tool will improve model building, population simulation and clinical trial simulation for MIPD.

3.2. Simulate MIPD for infliximab induction therapy in ulcerative colitis patients

Objective 2

By performing an in silico population simulation of infliximab precision dosing, the improvement on PD outcomes can be quantified.

Infliximab is an anti-TNFα inhibitor used in the treatment of ulcerative colitis. Unfortunately, therapeutic failure is a common occurrence, with as many as 40% of patients not responding. The drug is well tolerated, but drug cost makes achieving high exposure using high fixed doses prohibitively expensive. Based on an early drug concentration sample, underexposure to the drug may be avoided through dose increases. MIPD allows rational dose adaptation, ideally using this expensive biological drug more efficiently than administering a fixed one-size-fits-all dose. This was already demonstrated in

prospective trials for maintenance therapy by Negoescu et al.[42] Application of the techniques previously developed will allow us to quantify whether this promise hold up in silico for induction therapy, before an actual trial is conducted.

3.3. Simulate MIPD for tacrolimus in de novo kidney transplant recipients early post-transplant

Objective 3

By quantifying the improvement MIPD brings to tacrolimus target attainment, the clinical benefit vs required effort can be evaluated. A targeted clinical trial to show this clinical benefit can be efficiently designed.

Tacrolimus is an immunosuppressor used after solid organ transplant. Sufficient exposure is needed to limit organ rejection, while high exposure is nephrotoxic. This limited therapeutic window is further complicated by high PK inter-individual variability, with individual maintenance doses ranging from 3mg to 30mg per day to achieve the same target exposure. Therapeutic drug monitoring is required by the label to adapt the dose. The challenges of linking concentration target deviation to dose adjustment have been widely discussed,[43] and can be mitigated using a population PK model.

A retrospective dataset of 315 patients is available. Daily tacrolimus trough samples were routinely collected during the first 14 days post transplant. We will develop a population PK model to fit this data, ensuring the model is fit for use in MIPD. This model can then be used to predict the improvement between standard of care (SoC) and MIPD in the population. Finally, these predictions can then be used to design a clinical trial having appropriate power to detect this improvement.

---

[42] "Proactive Vs Reactive Therapeutic Drug Monitoring of Infliximab in Crohn's Disease".

[43] Wallemacq et al., "Opportunities to Optimize Tacrolimus Therapy in Solid Organ Transplantation"; Brunet et al., "Therapeutic Drug Monitoring of Tacrolimus-Personalized Therapy".

3.4. Build a tacrolimus MIPD software tool

Objective 4

If MIPD is predicted to deliver a clinically significant benefit to patients, a software tool will bring this technology in the hands of physicians.

MIPD for immunosuppressive drugs in solid organ transplant has been available since 1994,[44] at ever-increasing precision and ease of use.[45] Early on, clinicians filled out a form, faxed it to the clinical pharmacology department, and waited for dosing advice. Data was manually transcribed in digital format and dosing advice was computed using specialized FORTRAN programs, or (later) standard modeling tools. Software technology has evolved since, but having only a clinician in the loop remains a major challenge.

We envision an application with an easy-to-learn user interface, allowing physicians to enter patient covariates, treatment history and observed measures. The physician receives immediate feedback on the entered data, with the application showing the population typical prediction, the individually estimated prediction, and the prediction when implementing the dosing recommendation.

To optimize the physician workload, we further envision a connection with the electronic patient health records (EHR). The application should automatically receive EHR data, transform it appropriately and return a dosing recommendation to the EHR application. This should result in minimal disruption to the clinical workflow, increasing the tool adoption and limiting marginal costs.

[44] Anderson et al., "Evaluation of a Bayesian Approach to the Pharmacokinetic Interpretation of Cyclosporin Concentrations in Renal Allograft Recipients".

[45] Woillard et al., "Tacrolimus Exposure Prediction Using Machine Learning".

3.5. Transpose this approach to other compounds

Objective 5

The design of the software and approach should allow easy extension to other compounds.

Throughout the project, all methods and software are developed with extension to other compounds and disease areas in mind. By making this generic, we aim to build a platform for precision dosing of many compounds. This unmet need should attract other academic groups, increasing the number of developers/users. Such a critical mass[46] can then support an open, stable alternative to commercial efforts. Not only does this include publishing the software and methods, but also ensuring training and documentation is available.

[46] "Open source projects die because of an inability to acquire a critical mass of users", from 'Open Source Software Development as a Special Type of Academic Research', by Nikolai Bezroukov

Chapter 4 Methods for MIPD

This chapter explores how the classical modeling & simulation approach used in drug development should be adapted for MIPD development in five key areas: goodness of fit evaluation, covariate selection, residual error model evaluation, target selection, and population simulation. In 0, these methods are implemented. The clinical use cases in Chapter 1, Chapter 1, Chapter 7 and Chapter 1 use these methods to predict MIPD performance.

4.1. Goodness of fit evaluation

Evaluating whether a population model fits a clinical dataset is a well-studied problem. We evaluate whether the model predicts the observed values (DV vs PRED), whether residual errors are heteroscedastic and have no bias (IWRES vs TIME, CWRES vs TIME), and whether individual estimates reflect a normal distribution (NPDE for η distributions). We even verify whether a monte carlo simulation based on the model generates the same summary statistics as the clinical dataset (Visual Predictive Check and Numerical Predictive Check). However, what matters for precision dosing?

A model is useful for precision dosing if it accurately predicts future outcomes, with concentration target attainment most often used as a surrogate outcome. An exercise known as *prospective evaluation*[47] is required. We predict each observed value $i + 1$ using an EBE fit on all previous observed values 1 to i. Relevant summary statistics in the form of mean percentage prediction error (MPPE) and root mean squared error (RMSE) show bias and imprecision respectively. The full distribution of prediction errors can also be analyzed. This is not routinely done; existing literature models should be re-evaluated for their fitness for precision dosing.

[47] The technique was first implemented in the Perl-speaks-NONMEM tool suite, and the developers called it 'proseval'. In my humble opinion, the name is poorly chosen, as it can be confused with a 'prospective clinical trial'. We would propose 'individual predictive performance evaluation' as a more appropriate term.

The equations below show how predictive performance can be evaluated through prospective evaluation. $f(t, \eta)$ reflects the model prediction at time t for individual parameters η.

$$IPRED_1 = f(t_1, 0)$$
$$\text{for } i \text{ from } 1 \text{ to } n:$$
$$\quad \hat{\eta}_i = EBE(Observed_{1 \text{ to } i})$$
$$\quad IPRED_{i+1} = f(t_{i+1}, \hat{\eta}_i)$$
$$Err_i = Observed_i - IPRED_i$$
$$MPPE = \frac{1}{n}\sum_i \frac{Err_i}{Observed_i}$$
$$RMSE = \sqrt{\frac{1}{n}\sum_i Err_i^2}$$

Better yet, the relative residual error $\frac{Observed_i - IPRED_i}{Observed_i}$ can be directly related to error in predicted dose, and subsequently error in resulting concentration after dose adaptation. If we consider the case of a 1-compartment model with a single dose, then the trough concentration $CONC$ equals $CONC = D/V * e^{-k*\tau}$.

When we estimate individual pharmacokinetic parameters V and k, an error is introduced: $V' = V + \epsilon$ and $k' = k + \epsilon$. These parameters are used to predict the trough concentration $IPRED = D/V' * e^{-k'*\tau}$. Applying this equation to find the recommended dose D_{rec} to reach a target trough concentration TGT yields $TGT = D_{rec}/V' * e^{-k'*\tau}$. Applying this dose will actually result in a concentration of $C_{res} = D_{rec}/V * e^{-k*\tau}$. Using these equations, the prediction error $PE\%$ on the next trough concentration can be rewritten as follows:

$$PE\% = \frac{IPRED - CONC}{CONC}$$
$$= \frac{D * e^{-k'\tau}/V' - D * e^{-k\tau}/V}{D * e^{-k\tau}/V}$$
$$= \frac{e^{-k'\tau}/V' - e^{-k\tau}/V}{e^{-k\tau}/V}$$
$$= \frac{D_{rec} * e^{-k'\tau}/V' - D_{rec} * e^{-k\tau}/V}{D_{rec} * e^{-k\tau}/V}$$

$$= \frac{TGT - C_{res}}{C_{res}}$$

In this example, the prediction error on the next concentration $PE\%$ is equal to the concentration deviation from target after applying MIPD. The probability of target attainment for a given therapeutic window can thus be related to the prediction error. Assuming a target window between a and b, a target of $TGT = \frac{a+b}{2}$ maps to a prediction error between $\frac{TGT-b}{b}$ and $\frac{TGT-a}{a}$. This can be used as an evaluation of probability of target attainment after applying MIPD, and can thus be used to compare model performance.

For reference, $PE\%$ can also be translated to $C_{res} = \frac{TGT}{1+PE\%}$. The theoretical upper limit for probability of target attainment (assuming perfect individual parameter estimation) can be determined by the residual error, based on the probability that $\mathcal{N}(0, \sigma)$ is within the target of $\left[\frac{TGT-b}{b}, \frac{TGT-a}{a}\right]$.

4.2. Covariate selection

McDougall et al.[48] disruptively stated that the population model does not really matter for precision dosing. To make this point, they used a reference model and five intentionally misspecified models for voriconazole, and performed precision dosing with each model on a simulated dataset. Their results showed that, when sufficient concentration samples are available, only severe structural model misfit -e.g. removing a non-linear clearance route- resulted in different predicted doses. This is also illustrated in Figure **9**.

[48] "The Impact of Model-Misspecification on Model Based Personalised Dosing".

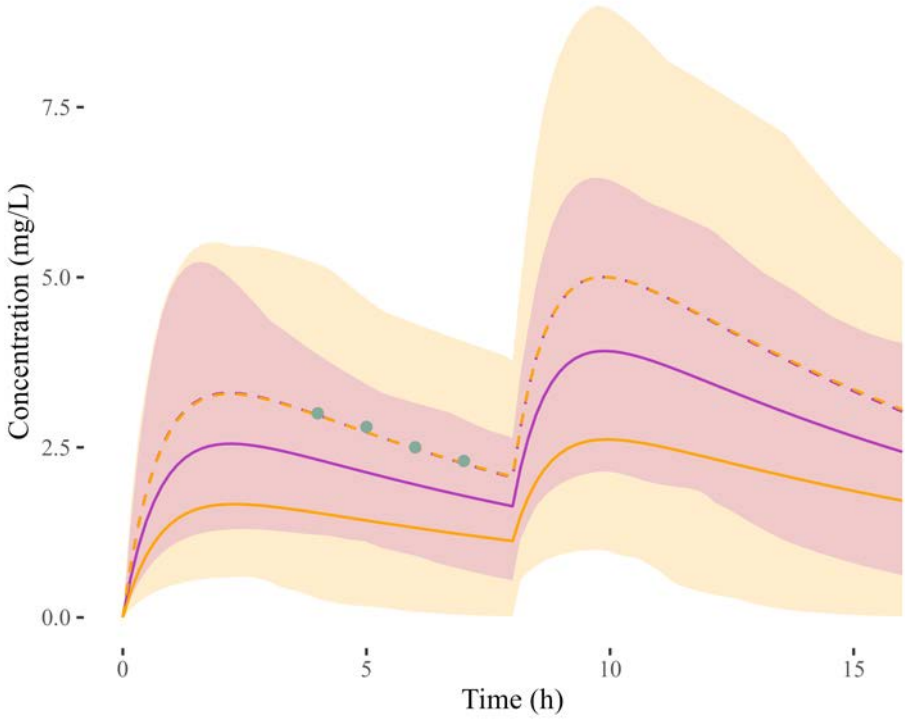


***Figure 9**: EBE fit (dotted line) and population prediction (median as solid line, 90% prediction interval area) for a patient with 45kg bodyweight. The purple model incorporates allometric scaling for weight (predicting typical higher concentrations for a patient weighing 45kg), while the orange model has no covariates. The estimated individual fit on 4 measured concentrations (blue points) is the same for both models.*

$$p(X_j|DV_j) \sim \Phi(\eta_j, \Omega) \times \prod_{i}^{n} \Phi\left(DV_{i,j} - Y_{i,j}, \Sigma\right)$$

A more nuanced view is to state that precision dosing has two separate phases. If we look at the equation for marginal likelihood above, it is composed of the population prior and the likelihood of observed values. If no observed values are available, the most likely estimate is the population prior; the patient is most likely a typical patient. The uncertainty on this estimate is as large as the population inter-individual variability ω. Therefore, when no observed values are available, individual precision dosing is equivalent to population fixed dosing. Any covariates that may help

to reduce unexplained IIV will also reduce η uncertainty, and improve dosing accuracy.

As more observed values become available, e.g. by measuring blood concentration after the first dose, the likelihood of observed values becomes more important when determining $\hat{\eta}$. Uncertainty shrinks, in line with the model residual error ($\sigma_{\hat{\eta}} \sim \frac{\sigma}{\sqrt{n}}$). McDougall et al. used a very rich sampling to demonstrate their point, while in reality there is a gradual change from prior estimate to posterior estimate as more information reflective of the individual becomes available. Without sufficient information, the EBE estimate $\hat{\eta}$ will always regress to the typical prior value.

Some argue that the estimated prior distribution hinders accurate predictive power. It may be appropriate to include uncertainty on priors, leading to a hierarchical bayesian estimation. Such estimation software exists in other domains, but has not yet been applied to pharmacometrics problems. Hughes and Keizer[49] showed that artificially inflating ω, a technique called *flattened priors*, also increases predictive performance. Neely et al.[50] argue that the use of a normal or lognormal prior for parameters is inappropriate altogether. They showed that the population distributions could be characterized by a non-parametric adaptive grid (NPAG) instead. Instead of defined statistical distributions with parameters (θ, ω), NPAG is a set of support points and associated weights approximating any probability distribution. As this method does not assume distribution shape, it can fit irregular distributions (e.g. bimodal) much better. NPAG should in theory be more appropriate for precision dosing, but has not been widely adopted.

A further pitfall in covariate search is to include time-varying covariates. Indeed, these may strongly explain between-occasion variability in parameters, improving population fit. However, this is of limited use in precision dosing, as the future time course of this covariate is not available. It is more appropriate to consider these covariates as measures -similar to drug

[49] "A Hybrid Machine Learning/Pharmacokinetic Approach Outperforms Maximum a Posteriori Bayesian Estimation by Selectively Flattening Model Priors".

[50] "Accurate Detection of Outliers and Subpopulations With Pmetrics, a Nonparametric and Parametric Pharmacometric Modeling and Simulation Package for R".

concentrations- in a joint model instead. Special care must be taken with covariates whose time course may be influenced by patient disease state, and these effects must be included in the model for appropriate dose simulation. Developing such a complex model may not be feasible; the base model without covariates and increased IIV may be more appropriately used instead.

4.3. Residual error models and MPC/MIPD

The key for EBE estimation is the balance between population prior and fitting the observed values. The latter is determined by both the nonlinear relationship between parameters and prediction, and by the residual error model. In other words, a difference between predicted value and observed value can be explained either by residual error ϵ, or by a change to η. There is an essential difference: a change to η will influence future predictions and therefore the recommended dose.

Whereas population modeling often considers the residual error model an afterthought, it is essential in precision dosing. Alihodzic et al.[51] showed that errors in recorded sampling time or dosing time lead to severe model mispredictions. It is worthwhile to refine residual error models to only incorporate assay error, and incorporate different errors (such as dosing timing error) instead. Unfortunately, this area is frought with numerical issues in estimating such models.

Another overlooked area is time dependence of residual error. In clinical reality, recent observations are more predictive of the immediate future. We discussed existing solutions in 1.4.2, yet we propose an alternative here: autocorrelation.

The existence of time-dependent residual error can be formally tested in a dataset using autocorrelation. Autocorrelation ρ_k is the correlation between a series and the same series shifted by k places, with k an integer value. $E[\rho_k]$ describes the expected (mean) autocorrelation, with $E[\rho_1]$ the mean autocorrelation for consecutive values, $E[\rho_2]$ the mean autocorrelation for values two days apart, etc.

[51] "Impact of Inaccurate Documentation of Sampling and Infusion Time in Model-Informed Precision Dosing".

In a classical empirical bayesian estimation (EBE), autocorrelation of the residual error (RE) for a given patient is driven in three ways. Firstly, assay error is assumed to be random, driving residual error autocorrelation of consecutive observations to $E[\rho_1] = 0$. Secondly, parameter regression to the mean results in a consistent residual error with $E[\rho_k] > 0$ for large values of k. Thirdly, model misspecification results in a consistent residual error of $E[\rho_k] > 0$ for small values of k; i.e. the model under- or overpredicts for a limited number of consecutive observations. By calculating the mean autocorrelation of the residual error for values $k = 1..n$, the amount and type of RE autocorrelation can be identified. A high autocorrelation that decreases for higher values of k points to model misspecification.

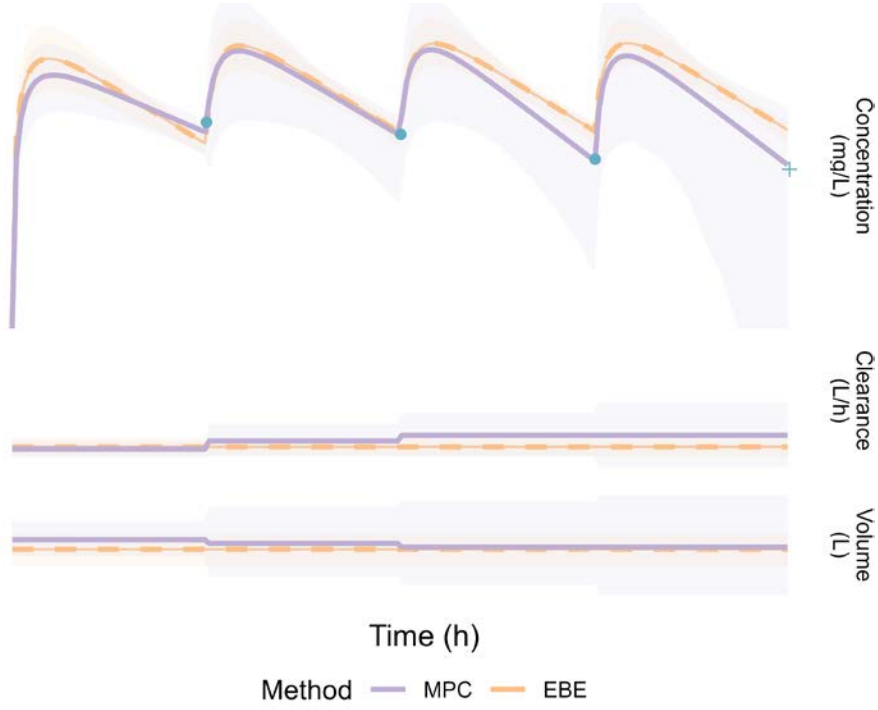


***Figure 10**: In the above example, we estimate the future concentration (+ symbol) for a patient with 3 blood samples (dots) who has just experienced renal failure. Drug clearance is almost non-existant, and restarts only gradually after day 2. Estimation using the EBE method (orange) fails to capture this aspect, assuming stable parameter values for clearance (CL) and volume (V1). The MPC method (purple) allows drifting estimates, assuming the previous estimated clearance will remain valid tomorrow. This also slightly overestimates concentrations, but the effect is much less dramatic.*

We developed an alternative way to integrate time dependence of residual error however (illustrated in Figure **10**). Intuitively, the individual parameter estimate $\hat{\eta}$ on day i can be considered a population prior for day $i+1$. A pragmatic way to adapt the model can then be implemented. Just as with a model including inter-occasion variability (IOV), we allow individual parameters to change every occasion. However, parameters move from their previous value, rather than from a mean IIV estimate as with classical EBE.

$$\begin{aligned} X_1 &= \theta_X * e^{\eta_{X,1}} \\ X_2 &= \theta_X * e^{\eta_{X,1}+\eta_{X,2}} \\ &\dots \\ X_i &= \theta_X * e^{\eta_{X,1}+\eta_{X,2}+\dots+\eta_{X,i}} \end{aligned}$$

When estimating an individual fit, only η_i for the current day i is estimated, with all past η fixed. If no information is available on day $i+1$, then η_{i+1} will be estimated at 0 and the estimate will be equal to day i. For pragmatic reasons, the original estimate of ω is used as a cost function when estimating each η_i.

It should be noted that MPC/MIPD, even though founded in pragmatism, is formally related to classical EBE. The technique approximates the individual objective function $OFV(\eta)$ by a new prior distribution around the mode of the previous estimate $\widehat{\eta'}$ on observations $1..i-1$:

$$OFV(\eta) = OFV_{pop}(\eta|\Theta,\Omega) + \sum_{0..i} O\,FV_{pred}(f(\eta)|Y,\Sigma)$$

$$OFV = OFV_{pop}(\eta|\Theta,\Omega) + \sum_{0..i-1} O\,FV_{pred}(f(\eta)|Y,\Sigma) + OFV_{pred}(f(\eta)|Y_i,\Sigma)$$

$$OFV \approxeq OFV_{pop}\left(\eta|\widehat{\eta'}, se\left(\widehat{\eta'}\right)\right) + OFV_{pred}(f(\eta)|Y_i,\Sigma)$$

In the above equation, we have approximated the objective function for an EBE of observations $1..i-1$ by a new individual prior around the mode $\widehat{\eta'}$. The Ω for this prior should be equal to the asymptotic standard error of the previous estimate, as this would approximate EBE as closely as possible. In this case however, we chose to keep the larger population IIV Ω instead, allowing the estimate to drift between occasions:

$$OFV_{MPCMIPD} = OFV_{pop}\left(\eta|\widehat{\eta'},\Omega\right) + OFV_{pred}(f(\eta)|Y_i,\theta,\Sigma)$$

4.4. Target selection

When dosing a drug, we aim for optimal safety and efficacy. Through empiric studies, we relate acceptable safety and efficacy to an acceptable concentration range. When a one-dose-fits-all approach fails to put all patients in this acceptable range, therapeutic drug monitoring may be performed. Being outside of the target range is bad, and results in a dose adaptation.

This method is not appropriate. It is a simplification, performed for pragmatic reasons. Holford, Ma, and Metz[52] argued that we should discontinue the use of a therapeutic window for dosing decisions. There is no *range* of optimal exposure, there is just a single *optimal exposure* associated with optimal outcomes. Although intrinsic variability in clinical outcome may be far greater than the difference in average outcomes across the therapeutic window, thwarting the identification of this optimal exposure, this is ultimately irrelevant. The principle stands: doses should be adapted to target optimal exposure.

The work done by Holford et al is important, yet only proposes a small incremental step. Precision dosing should target optimal safety and efficacy, which is subject to high inter-individual PD variability for a given target exposure. Ideally, we may measure relevant biomarkers directly related to safety and efficacy. These are then related to a clinical utility score. We reformulate the precision dosing problem as follows: find the optimal dose that maximizes individual clinical utility. Let us look at an example.

```
f() = {
  KA = TV_KA
  CL = TV_CL * exp(ETA_CL)
  V = TV_V * exp(ETA_V)
  d/dt(A0) = -KA*A0
  d/dt(A1) = KA*A0 - CL/V*A1
  CONC = A1 / V

  EC50 = TV_EC50 * exp(ETA_EC50)
  Kin = TV_Kin * exp(ETA_Kin)
  Kout = TV_Kout * CONC / (EC50 + CONC)
```

[52] "TDM Is Dead. Long Live TCI!"

```
  d/dt(SLD) = Kin - Kout * SLD

  EC502 = TV_EC502 * exp(ETA_EC502)
  Kin2 = TV_Kin2 * exp(ETA_Kin2) * CONC / (EC502 + CONC)
  Kout2 = TV_Kout2
  d/dt(ALT) = Kin2 - Kout2 * ALT

  CU = W1 * (1/SLD) - W2 * (ALT > 150)
}
```

An example model of a chemotherapy drug incorporates four interlinked submodels. A first model predicts drug concentrations using a 1-compartment model with oral absorption. IIV is applied on CL and V. A second model predicts the sum of largest tumor diameters (reflecting current cancer burden) by an indirect response model. Drug concentrations increase tumor cell death, modeled as a drug effect on $Kout$. A third model predicts ALT liver enzymes, with drug concentrations causing an increased production of ALT. Finally, a fourth model predicts clinical utility CU by both SLD and ALT levels. These surrogate markers are weighed using $W1 = 10$ and $W2 = 0.1$ (respectively), and weights can be determined by physicians or even patients according to individual preference.

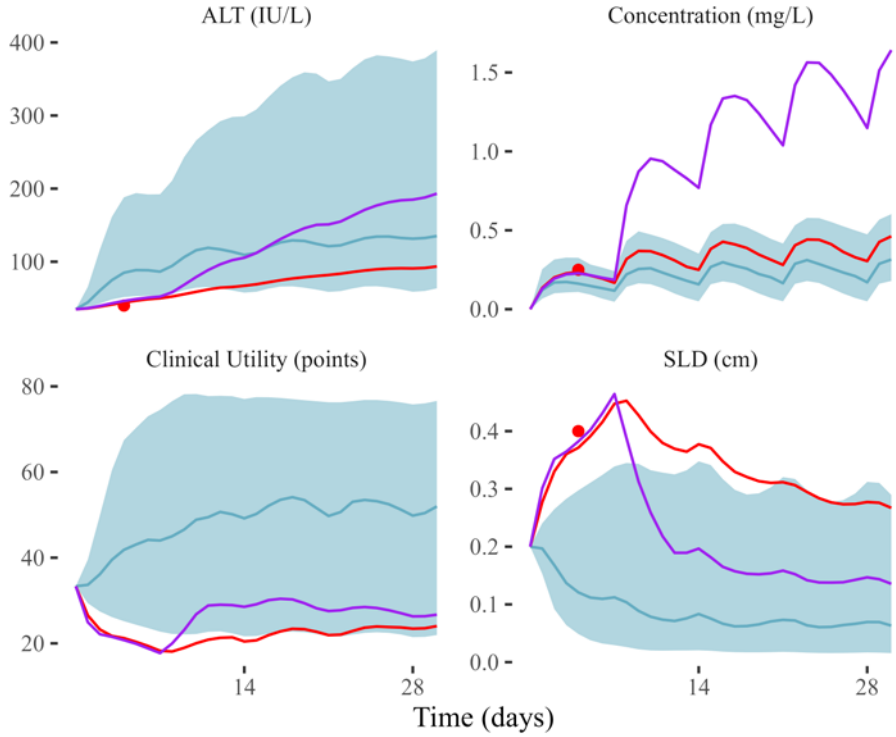


***Figure 11**: Population typical prediction (blue line) and prediction interval (blue area) for ALT, drug concentration, sum of largest tumor diameter (SLD) and clinical utility (CU) at the administered dose of 1mg. The measured ALT, concentration and sum of largest diameter (red points) inform the EBE fit (red line). If 1mg is continued, SLD is predicted to decrease only slowly, ALT will remain low, and CU is poor. Adapting to 3mg (purple line) is predicted to result in increased efficacy (low SLD), acceptable liver toxicity (low ALT), and better overall clinical utility (CU).*

When performing empirical bayesian estimation to determine individual variability on these interlinked models (η_{CL}, η_V, η_{EC50}, η_{Kin}, $\eta_{EC50,2}$, and $\eta_{Kin,2}$), all available measurements can be used: concentration measurements in blood, sum of largest tumour diameter from CT scans, or ALT levels in blood. This principle is demonstrated in Figure **11**. Future dose recommendations can be made targeting optimal predicted clinical utility. As stated in the Introduction chapter, a slow transition to this model is already ongoing, with increasing pressure from regulators to explore patient

exposure-response of compounds, rather than simply establishing empiric "good" one-size-fits-all dosing.

4.5. Population simulation for predicting MIPD

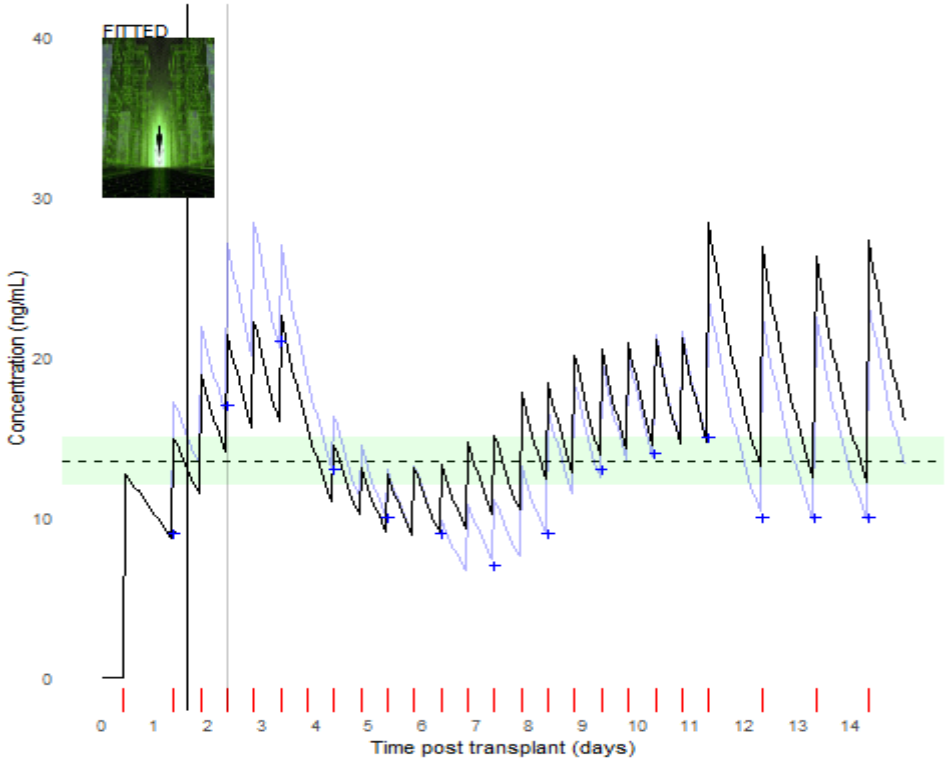


***Figure 12**: Animation of how MIPD is simulated on retrospective data for a single patient. The algorithm alternates between three steps (shown in top left). It first simulates ("MEASURED") the next concentration using the full fit (lightblue line) and the proposed dosing regimen, resulting in a black dot. It then fits this concentration ("FITTED"). The black line shows the computer prediction for tomorrow. The algorithm now adapts future doses so all predicted trough concentrations hit the target ("ADAPTED", target window in green). It then advances the current time (black vertical line) and performs a new iteration. Red markings on X axis show the time of true observations, blue crosses show the predicted concentrations using the currently proposed treatment regimen. The black points show the simulated concentrations for precision dosing. Available online at http://bit.ly/3t7pl00*

In section 2.6, we derived a relationship between prediction error and concentration error after precision dosing. This relationship does not hold for more complex precision dosing, such as discussed in section 4.4. In these cases, simulation is required.

```
INITIALIZATION:
   observed = []
   regimen = [ Loading Dose, Future Planned Doses]
   eta_true = monte carlo sample from population prior
FOR EACH SAMPLING TIMEPOINT t_i:
 SIMULATE Y_i^'≔ f(eta_true,t_i ) + epsilon
   add (t_i,Y_i) to observed
   ESTIMATE eta using MPC/MIPD
   OPTIMIZE the treatment regimen using eta
```

In the above algorithm, we perform an iterative loop of simulate/estimate/optimize (see also Figure **12**) for a hypothetical patient. We first simulate a future concentration using the true parameters η_{true}, the current treatment regimen, and a sampled residual error ϵ. This simulated concentration is used to estimate η. This approach works well to explore non-linear properties of precision dosing, such as model fitness for precision dosing, or blood sampling scheme performance.

Alternatively, we can use a retrospective dataset for this simulation. η_{true} is then equal to the posthoc estimate on all real observed data for that patient. Instead of sampling ϵ, it can be reused from the posthoc estimate as well. Assuming the model predicts concentration on adapted doses appropriately (i.e. the model interpolates different doses correctly), resulting concentrations will be more realistic, as the residual error is reused. Fewer assumptions are required in this approach.

Chapter 5 Tdmore - a framework for model-informed precision dosing

In this chapter, we describe the design and functionality of the `tdmore` software, a flexible and adaptable framework for model evaluation, performance prediction, and bedside execution of precision dosing. This work was funded by FWO TBM grant T003117N (OCT-2017 to OCT-2020). Part of the software is available from github.com/tdmore-dev/tdmore, and has been reproduced below in simplified form.

Loosely based on:
Ruben Faelens, Nicolas Luyckx, Quentin Leirens, Dirk Kuypers, Pieter Annaert (2020). Building model-informed precision dosing software using R: blueprint for a state-of-the-art development process. *Presented at PAGE 2021 conference as a scientific poster*

5.1. Introduction

> The proof of the pudding is in the eating, but you cannot recoup the eggs and milk from a bad pudding.

Software for precision dosing is not new. A study by Kantasiripitak et al.[53] identified 28 software tools for MIPD, with the aim of benchmarking these tools for use in a clinical setting. 11 tools were not actively maintained, and 2 tools did not use a Bayesian approach. One tool did not include a user interface, 2 tools only supported one drug, and 2 software providers declined participation. This yielded 10 software tools that were benchmarked on evaluation criteria "related to (i) user-friendliness and utilization, (ii) user support, (iii) computational aspects, (iv) population models, (v) quality and validation, (vi) output generation, (vii) privacy and data security, and (viii) cost". This study focused on the use of software tools by clinicians. Tools generally performed well, with scores ranging from 7.2 ± 2.1 to 8.5 ± 1.8 (out of max 10 points). We can conclude that software tools for precision dosing by clinicians is a well-served market segment.

We revisited the broader selection of 15 Bayesian tools with active maintenance, and evaluated these based on the aforementioned objectives (see Chapter 1). Criteria can be grouped in 3 categories. First, we evaluated adaptability for pharmacometricians: how easy can a pharmacometrician define their own PKPD model, associated PKPD targets and search domain? Second, we evaluated user interface flexibility: how easy can the interface be adapted to fit within an existing clinical workflow? Third, we evaluated simulation capabilities: can the tool be used to predict MIPD performance?

[53] "Software Tools for Model-Informed Precision Dosing".

***Table 1**: Evaluation of MIPD software on adaptability, flexibility and simulation.*

Name	Custom model	Custom target function	Custom workflow	Developm ent support	Simulation	Source code
TDM for R[U]	No*	No*	No*	Sparse	No*	Yes, under GPL2/GPL 3
DosOpt[U]	No	No	No	Unknown	No	No
myPKFit[P]	No	No	No	N/A	No	No
iDose[C]	No	No	No[+]	N/A	No	No
RxKinetics[C]	Custom parameter s	No	No	N/A	No	No
AutoKinetics[U]	No[+]	No[+]	N/A	N/A	No	No
BestDose[U]	Yes	Yes	No	Unknown	Yes	No
DoseMeRx[C]	No[+]	No[+]	No[+]	N/A	No[+]	No
ID-ODS[C]	No	No	No	N/A	No	No
InsightRx Nova[C]	No[+]	No[+]	No[+]	N/A	No[+]	No
MwPharm++[C]	Yes	Yes	No	N/A	No	No
NextDose[C/U]	No	No	No	N/A	No	No
PrecisePK[C]	No	No	No[+]	N/A	No	No
TDMx[U]	Custom parameter s	No	No	N/A	No	No
Tucuxi[C]	No	No	No	N/A	No	No

U: Developed by university, C: Commercially developed, P: Developed by pharma industry

*: As the source code is available, this can be implemented by the user.

+: By manufacturer, at their discretion

Results are shown in Table **1**. Only one software tool made the source code available. 4 manufacturers advertised their willingness to adapt the software to an existing clinical workflow. 2 manufacturers offered simulation services

to predict MIPD performance, and one more manufacturer offered the mathematical engine to allow users to perform MIPD performance predictions. While these 15 software tools scored well for use by physicians, there is clearly an unmet need in adaptability, custom development, and simulation.

Further demonstrating this unmet need is the approach InsightRx proposes[54] to implement MIPD. Rather than selling a commercial off-the-shelf software with little user support, the company proposes a multi-step process. First, the proposed PK/PD model is qualified for predictive performance on the institution's own retrospective data. Next, the company allows integration of their MIPD software into EHR systems, simplifying the clinical workflow. Finally, they offer a data dashboard for hospital administrators to analyze MIPD performance, measured as *"% of patients achieving therapeutic range, time to therapeutic range, and time within therapeutic range"*.

The evidence-based process outlined above is certainly appropriate for drugs with established MIPD benefits, and *a posteriori* evaluation and continuous learning of MIPD implementation is commendable. **However, we stress that this happens *after* the initial investment into MIPD software, based on prior belief that MIPD brings benefits.** This is not appropriate for new drugs, where such evidence is lacking.

Furthermore, while this approach has a solid scientific base, the company aims to leverage this as a competitive advantage to increase profits, carefully protecting its intellectual property. Based on this approach, InsightRx raised a total investment of $12.8M. Competitor DoseMe also offers tailored solutions, and raised $3.1M total. Once a hospital has decided to work with a commercial MIPD company, they are locked in[55] to that customized solution. A comfortable position for commercial companies aiming to rake in license fees for years to come, yet a less fortunate position for the spending of public health funds. A growing voice[56] suggest open-source medical devices as an

[54] Based on publicly available marketing material, as described on https://www.insight-rx.com/, consulted 17-NOV-2021

[55] Ven, Verelst, and Mannaert, "Should You Adopt Open Source Software?"

[56] Ahluwalia et al., "Towards Open Source Medical Devices"; Winter et al., "Open Source Medical Devices for Innovation, Education and Global Health".

alternative to promote sustainable growth and innovation, while democratizing access to these technologies.

In summary, commercial efforts focus on user-friendly software for physicians, with little support to incorporate custom models, predict MIPD performance before implementation, or adapt software to specific clinical workflows, except as a means to increase customer retention. The `tdmore` software aims to fill this unmet need by providing an open-source framework for precision dosing, and accompanying roadmap for MIPD evaluation and implementation.

5.2. The road to MIPD

Conceptually, precision dosing is founded on three basic elements. The first is a relevant target: a goal for the intervention. Ideally, we target optimal clinical outcome. Often, the target is a drug exposure or biomarker that has been empirically shown to correlate with favorable clinical outcomes. This correlation is demonstrated during drug development and further refined during clinical use. The second element is a way to relate the intervention to this target by way of a population PK/PD model. This model predicts an interval of possible outcomes in the population for any given intervention. For many drugs, a single dosing regimen can be found where this interval is sufficiently narrow that all patients will reach favorable outcomes: safe and effective drug dosing. For drugs with a narrow therapeutic interval however, all possible dosing regimen result in a too wide prediction interval: effective but toxic for some, safe but ineffective for others. Here, we introduce the third element: an individual measurement -drug concentration, biomarker, or effect- that allows to reduce the individual patient's prediction interval and identify a safe and effective dosing regimen.

A roadmap towards MIPD lies in applying these three elements in a structured way:

1. Through characterization of drug exposure-response, determine a target that is associated with favorable clinical outcomes.
2. Build a population pharmacometric model that relates intervention to the target.
3. Simulate whether a safe and effective fixed dosing exists.
4. Ensure the population model is appropriate for precision dosing, adapt if necessary.

5. Simulate how incorporating covariates or individual measurements reduces variability and improves outcomes. Decide whether this improvement is clinically meaningful.
6. Optionally, use these simulation results to design a clinical trial.
7. Build a precision dosing software useable by physicians
8. Integrate this software in a clinical workflow with minimal disruption.

The tdmore framework was built to support this roadmap. A first package `tdmore` was designed to define the pharmacometric model, perform empirical bayesian estimation and dose adaptation. This mathematical engine can be applied to an individual patient, supporting the real-life application of MIPD. The framework also allows to use a retrospective dataset or virtual simulated population. We can characterize model predictive performance, as well as dose adaptation performance.

If these simulated results show MIPD is worthwhile, the model can be easily incorporated in a software tool useable by physicians. To this end, the `shinytdmore` framework was constructed. It is a collection of user interface elements that can be assembled into a user interface. The inherent flexibility allows a developer to adapt to any clinical workflow, including or removing visual elements as required.

This roadmap was followed in the development of precision dosing for tacrolimus in kidney transplant recipients. The finalized software tool was composed of the mathematical engine, assembled user interface components, and appropriate pharmacometric model. The package was complemented by a method for the electronic health record (EHR) system to push data and relay back dosing recommendations.

5.3. Mathematical engine

At its core, the mathematical engine predicts a structural model. The model can be specified either as an RxODE[57] model, an algebraic function, a

[57] Fidler et al., *RxODE*.

deSolve[58]-compatible function, or included in an nlmixr[59] model. `tdmore` includes a `model_predict()` function that translates dosing regimen and inter-individual parameters into the appropriate call for each model and returns predictions in a standard format.

```
m1 <- RxODE::RxODE("
Ka=0.8
CL = 0.3 * (WT/70)^0.75 * exp(ECL)
V = 3 * (WT/70) * exp(EV)
Ke = CL / V
d/dt(A0) = -Ka*A0
d/dt(A1) = Ka*A0 -Ke*A1
CONC = A1/V
")
prediction <- tdmore:::model_predict(
  m1,
  times=seq(0, 12),
  regimen=tibble(TIME=0, AMT=30),
  parameters=c(ECL=0.2, EV=0.1),
  covariates=c(WT=60)
)
```

This is then passed to a `tdmore()` function and supplemented with the statistical model: inter-individual variability, inter-occasion variability (if applicable), and the residual error model. Alternatively, this information can be read directly from an `nlmixr` population model. Several error models are provided, and the user may optionally provide their own function to calculate log-likelihood. Metadata can also be added that aid in plotting, dose finding, or constructing the user interface.

```
model <- m1 %>%
  tdmore(omega=c(ECL=0.6, EV=0.4),
         res_var=list(errorModel("CONC", prop=0.1))) %>%
  metadata(formulation("HardCaps", unit="mg", dosing_interva
l=8,
                 default_value=10, round_function=round)) %>%
```

[58] Soetaert, Petzoldt, and Setzer, "Solving Differential Equations in *r*".

[59] Fidler et al., "Nonlinear Mixed-Effects Model Development and Simulation Using Nlmixr and Related R Open-Source Packages".

```
  metadata(formulation("ExtendedRelease", unit="mg", dosing_
interval=24,
                 default_value=50, round_function=round)) %>%
  metadata(covariate("WT", "Weight", "kg", min=40, max=120))
%>%
  metadata(output("CONC", "Concentration", "mg/L")) %>%
  metadata(target(min=3.5, max=4))
```

This model can then be used for empirical bayesian estimation. To this effect, the marginal likelihood function was implemented in tdmore as `pop_ll()` and `pred_ll()`. As this function is called frequently, the Cholesky[60] decomposition of the omega matrix is used for increased computational speed.

```
pop_ll <- function(par, model) {
  pdf <- mvnfast::dmvn(X = par, mu = 0, sigma = model$omega,
            log = TRUE, isChol = TRUE)
  sum(pdf)
}
pred_ll <- function(par, model, observed, regimen, covariate
s) {
  ipred <- stats::predict(model, newdata=observed, regimen,
par, covariates)
  ll <- 0
  for(var in model$res_var)
    ll <- ll + var$ll(ipred[[var$var]], observed[[var$var]])
  ll
}
ll <- function(...) {
  pop_ll(...) + pred_ll(...)
}
```

The log likelihood function can be used to explore the log-likelihood profile in an interactive way through a `shiny`[61] gadget. It is also used to find the maximum a posteriori estimate through the `estimate()` function. Note that

[60] Tanabe and Sagae, "An Exact Cholesky Decomposition and the Generalized Inverse of the Variance–Covariance Matrix of the Multinomial Distribution, with Applications".

[61] The Shiny framework allows R developers to easily build user interfaces.

a `model_prepare()` call is used to allow the structural model to set up caching, transform datastructures and perform all other work that can be done in advance. Subsequent calls to `ll()` (and thus `model_predict()`) will only entail modified `eta` estimates, and should be calculated as efficiently as possible.

For estimation of individual parameters through numerical optimization of marginal likelihood, the `optimr` package offers several possible algorithms. This flexibility is important for models with irregularly shaped log-likelihood profiles, such as when estimating inter-individual variability on absorption lag time. It can also use different starting points to ensure local maxima are avoided. All of this functionality is made available in tdmore for added robustness.

```
estimate <- function(model, observed, regimen, covariates, .
..) {
  model_prepare(model, observed, regimen, covariates)
  ofv <- function(par) {
    -2 * ll(par, model, observed, regimen, covariates)
  }
  res <- optimr::optim(ofv, init=0, hessian=TRUE, ...)
  varcov <- solve(res$hessian / 2)

  tdmorefit(model, observed, regimen, covariates, res$par, v
arcov)
}
```

Once a `tdmorefit` object is obtained, this can then be used to make future individual predictions. The function samples uncertainty from the variance-covariance matrix. As the hessian is not always positive semidefinite, markov-chain monte carlo sampling[62] on the marginal log likelihood is available as an alternative. A convenient `plot()` function is available, showing population prediction and individual prediction, as well as debugging functions to show the log-likelihood profile.

```
observed <- data.frame(TIME=c(7, 9), CONC=c(3.4, 14.1) )
regimen <- data.frame(TIME=c(0,8,16), AMT=15)
covariates <- c(WT=65)
fit <- tdmore:::estimate(model, observed, regimen, covariate
```

[62] Martin, Quinn, and Park, "**MCMCpack**".

```
s)
plot(fit, newdata=seq(0, 24, by=0.1))
```

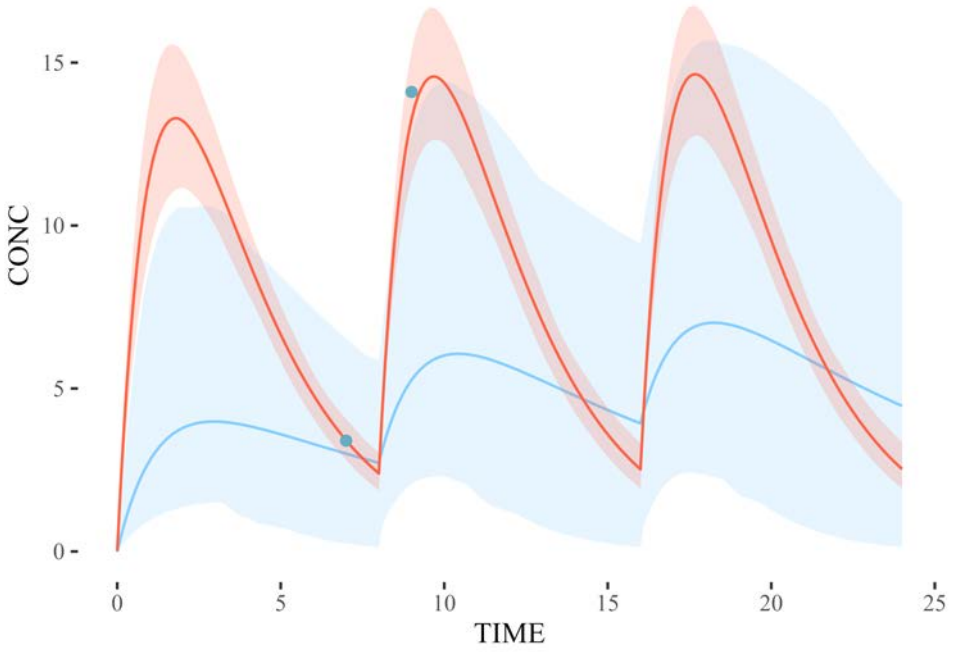


***Figure 13**: Concentration-time curve of typical value prediction (blue line) and EBE fit (red line and 90% confidence interval area) on observed points (grey points).*

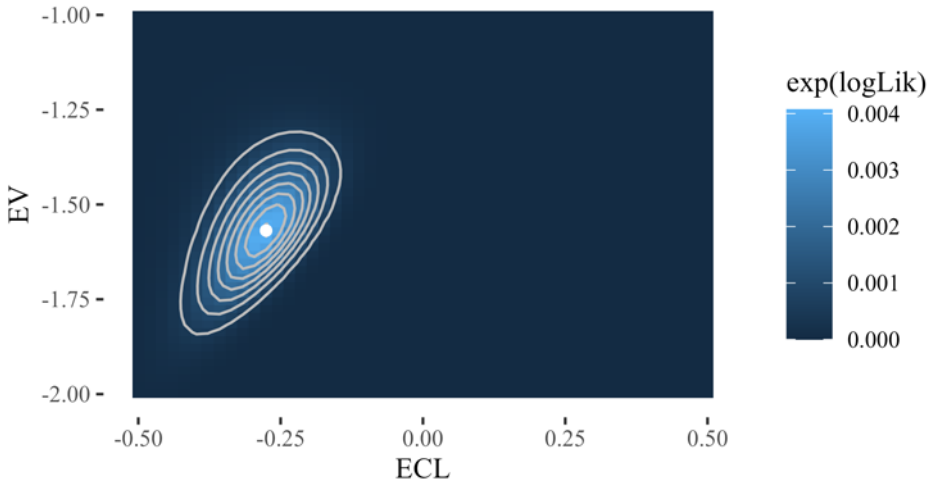


***Figure 14**: Contour plot of log-likelihood for combinations of η_V and η_{CL}. The plot visualizes uncertainty and correlation between the individual estimates.*

Finally, a `findDose()` function is used to optimize the dosing regimen. This uses a root finding function to identify what dose should be used on a given row of the treatment regimen to hit a pre-defined target perfectly, meaning $target - prediction = 0$. The `findDoses()` function extends on this to optimize multiple doses piece by piece, selecting the appropriate target per administration. Trough times can be detected automatically based on time of treatment + dosing interval, substituting this for actual planned treatments if appropriate. This finally results in a plot of individual dosing recommendation, showing population prediction (blue), individual fit with current regimen (red) and individual fit with recommended regimen (green). Measures used for fitting are represented as points, and the target concentration is represented as a target reticle.

```
findDose <- function(fit, doseRows, interval=c(0,9999), targ
et) {
  rootFunction <- function(AMT) {
```

```
    newRegimen <- fit$regimen
    newRegimen$AMT[doseRows] <- AMT
    pred <- predict(fit, newdata=target, regimen=newRegimen)
    pred - target
  }
  runUniroot(rootFunction, interval)
}

recommendation <- findDose(fit, doseRows=3, target=data.fram
e(TIME=24, CONC=8))
print(recommendation)

## Recommendation:
## A dose of ` 49.28681 ` will hit the requested target of
##  TIME CONC
##    24    8

plot(fit) + autolayer(recommendation) + coord_cartesian(xlim
=c(0, 25))
```

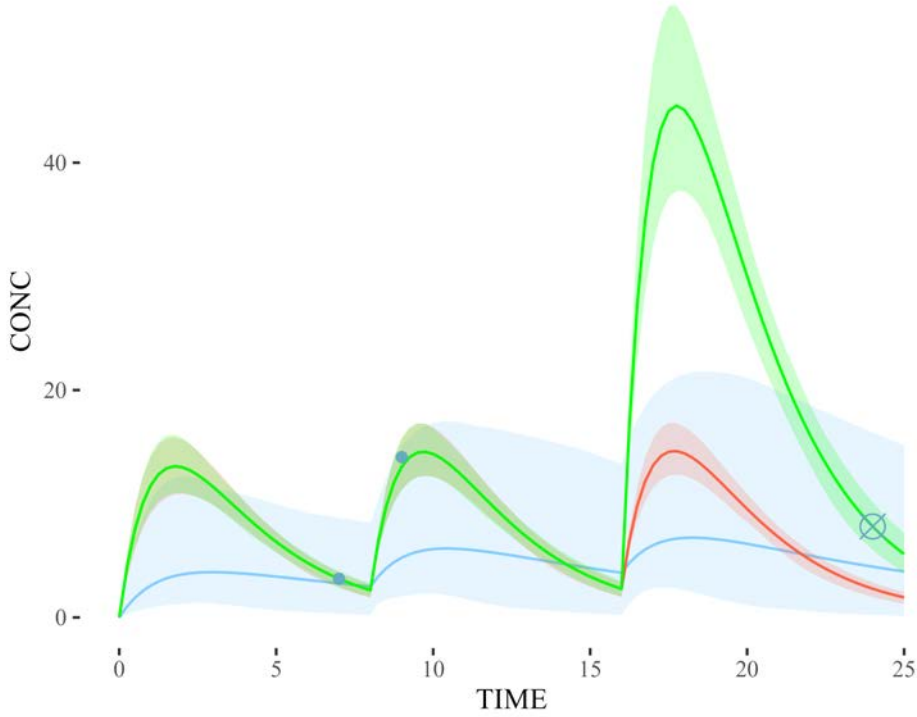


***Figure 15**: Concentration-time curve of typical value prediction (blue line), EBE fit (red line and 90% confidence interval area) on observed points (grey points), and prediction for recommended dose (green line and 90% confidence interval, dosing target as grey target reticule).*

5.4. Methods: a priori simulation

5.4.1. Simulation dataset

To predict MIPD performance, `tdmore` is designed around a step-wise evaluation of model and dose adaptation rules. First, a simulation dataset needs to be created. Ideally, a retrospective dataset representing Standard of Care is used. Based on a general patient assessment, physicians may sometimes adapt doses, and this *in cerebro* dose adaptation is difficult to replicate in silico. Lacking such a dataset, we may still simulate through monte carlo sampling from an a priori distribution. To do this, the `as.population()` function is used to provide a tdmore fit without any observations, yielding the typical values as estimate and a priori inter-

individual variability as parameter uncertainty. We can predict from this fit, yielding a population simulation. We sample with uncertainty (`se.fit=TRUE`), keeping all samples rather than summarizing (`level=NA`). Afterwards, the `model.frame()` function is used to add sampled residual error.

```
N <- 64 #for performance reasons
regimen <- data.frame(TIME=0, AMT=15, II=8, ADDL=8)
covariates <- c(WT=70)
pop <- as.population(model, regimen=regimen, covariates=cova
riates)
simulated <- predict(pop, newdata=c(7,10,15,48), se.fit=TRUE
, mc.maxpts=N, level=NA)
simulatedRe <- simulated %>%
  tdmore:::model.frame.tdmore(model, ., se=TRUE, level=NA)
trueParameters <- simulatedRe %>% rename(ID=sample) %>% grou
p_by(ID) %>%
  summarize(
    fit = list(tdmore:::tdmorefit(model,
        regimen=regimen,
        covariates=covariates,
        res=c(ECL=ECL[1],EV=EV[1])))
  )
observed <-  simulatedRe %>%
  filter(TIME %in% c(7,10,15)) %>%
  transmute(ID=sample, TIME, CONC)
db <- dataTibble(object=model, regimen=regimen,
                 observed=observed, covariates=covariates)
```

In this example, we use the previously defined model to simulate a dosing regimen of 15mg every 8h, with observations at 7h, 10h, 15h and 48h after the first dose.

5.4.2. Posthoc

As a first evaluation step, a fit on all available data is performed. This is known as a *posthoc* fit. This exercise is performed as part of standard population modeling. We will therefore not discuss it further here.

```
posthocfit <- posthoc(db)
```

5.4.3. Prospective evaluation

To evaluate predictive performance, a prospective evaluation is performed (see 4.1 for more details). The resulting graph shows the prediction using 0 (solid line, typical value prediction), 1 (dotted line), 2 (striped line) and 3 observations.

```
prosevalfit <- proseval(db)
```

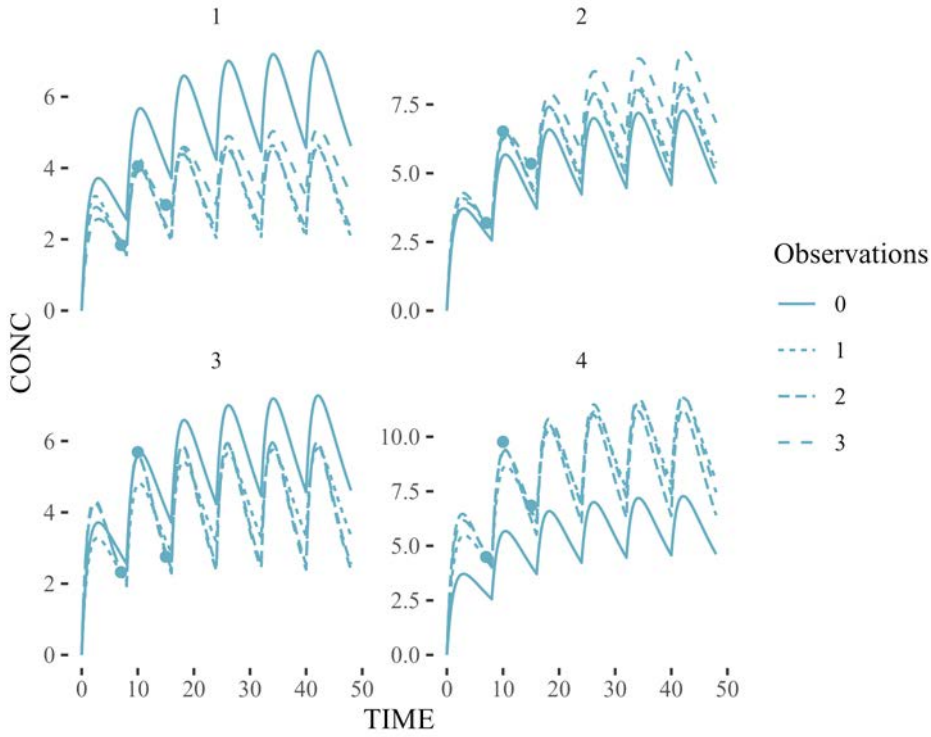


***Figure 16**: Predicted concentration-time curves for an EBE fit (line shapes, cf. figure legend) using 0, 1, 2 or 3 observations (points) for 4 virtual patients (seperate panels).*

As described previously, prospective evaluation results can directly translate into MIPD accuracy, and are therefore highly relevant. In this case, we can simulate how accuracy to predict steady-state trough (represented by the observation at 48h post-start) may improve with more observations. Predictive performance to predict concentration at 48h post treatment initiation is shown in the figure below, as box-plots of relative prediction

error versus number of observations. Boxplots show median (line), inter-quartile range (box), 1.5 times inter-quartile range (whiskers, roughly corresponding to 95% prediction interval) and outliers beyond whiskers. Only sampling at 7h, 10h and 15h post-initiation is predicted to reduce the 95% PI to acceptable ranges.

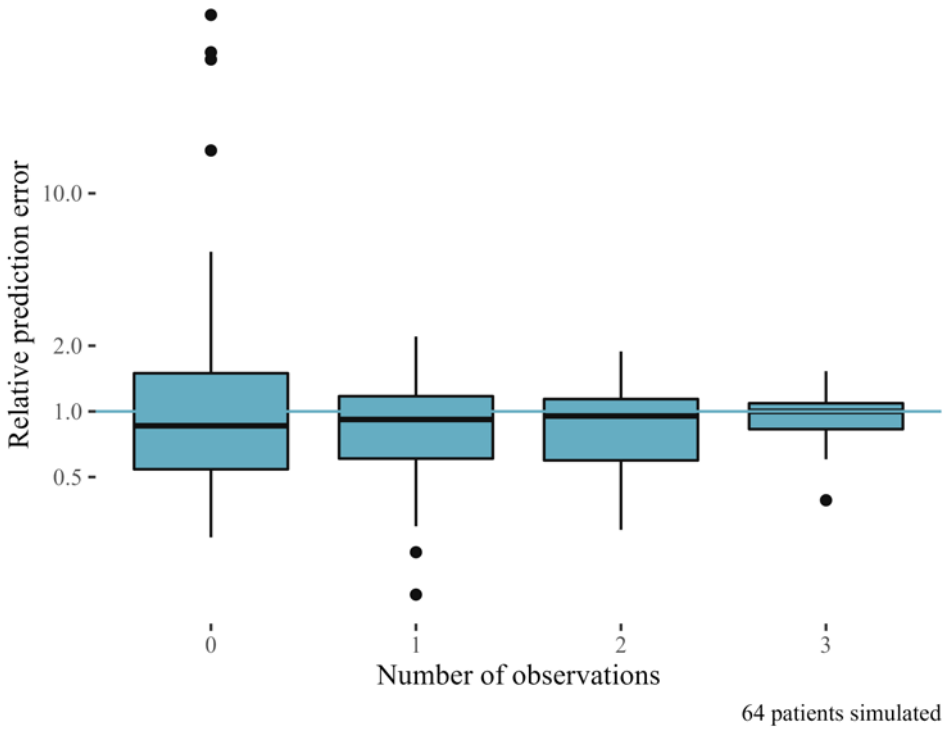


***Figure 17**: Relative prediction error of concentration at 48h (boxplots) for 64 simulated patients when using 0, 1, 2 or 3 observations in the EBE fit. A low prediction error leads to more accurate dosing recommendations.*

5.4.4. Dose simulation

Finally, we can predict how MIPD may adapt dosing (see 4.5 for more details). The `doseSimulation()` routine takes care of the three defined steps in iteration: simulate, estimate and optimize. To simulate the next measured concentration under a modified treatment regimen, we use the individual parameters originally sampled. If a retrospective dataset is used, we may use the most likely parameters as calculated by the posthoc fit.

Residual error is added, either sampled (in case of a virtual population) or reused from the posthoc fit (in case of a retrospective dataset). To estimate, the standard routines from `tdmore` are used. To optimize, a custom function is provided.

In this example, we investigate two dose adaptation methods. Both methods simply measure at two different timepoints, and optimize the treatment regimen to target a trough between 3.5 and 4 mg/L. Method A measures at 7h and 10h after the first dose, method B measures 7h and 15h after the first dose. We evaluate target attainment at 48h after start of treatment.

```
regimen <- data.frame(TIME=seq(0, 47, by=8), AMT=15, FORM="H
ardCaps")
simulationDb <- trueParameters %>%
  filter(ID <= 32) %>% #for performance
  select(ID, fit) %>%
  mutate(regimen=list(regimen),
         covariates=list(covariates),
         object=list(model))


measureAtSpecificTimes <- function(fit, regimen, truth) {
  now <- max(c(0,fit$observed$TIME))
  regimen$FIX <- !(regimen$TIME > now)
  rec <- findDoses(fit, regimen=regimen)
  i <- which( measureTimes > now)[1]
  nextTime <- measureTimes[i]
  list(nextTime=nextTime, regimen=rec$regimen)
}

measureTimes <- c(7,10)
result1 <- simulationDb %>%
  doseSimulation(optimize=measureAtSpecificTimes) %>%
  mutate(method="A. At 7h and 10h")
measureTimes <- c(7,15)
result2 <- simulationDb %>%
  doseSimulation(optimize=measureAtSpecificTimes) %>%
  mutate(method="B. At 7h and 15h")
result <- bind_rows(result1, result2)
```

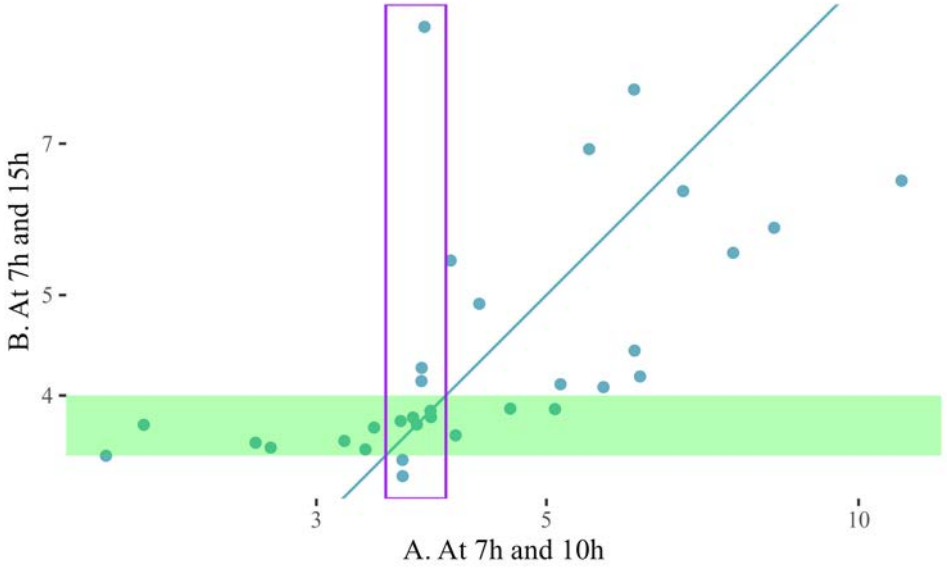


***Figure 18**: Predicted steady-state trough concentration after dose adaptation for sampling strategy A (x axis) and sampling strategy B (y axis) for 32 simulated patients. Patients close to the identity line (grey) have similar benefit in both sampling strategies. Patients in green box (but not purple) benefit more from strategy B, patients in purple box (but not green) benefit more from strategy A.*

A graphical comparison of predicted steady-state trough samples is shown in the figure above, comparing dose adaptation method A (x axis) to B (y axis) in the same patient, represented as individual points. The target range of 3.5 to 4 mg/L is shown as green and purple rectangles. Many patients are close to the identity line; methods A and B show similar results in these patients. However, method B severely under- and overdoses some patients, while method A puts these patients in target. Therefore, method A can be regarded as superior.

5.5. User interface

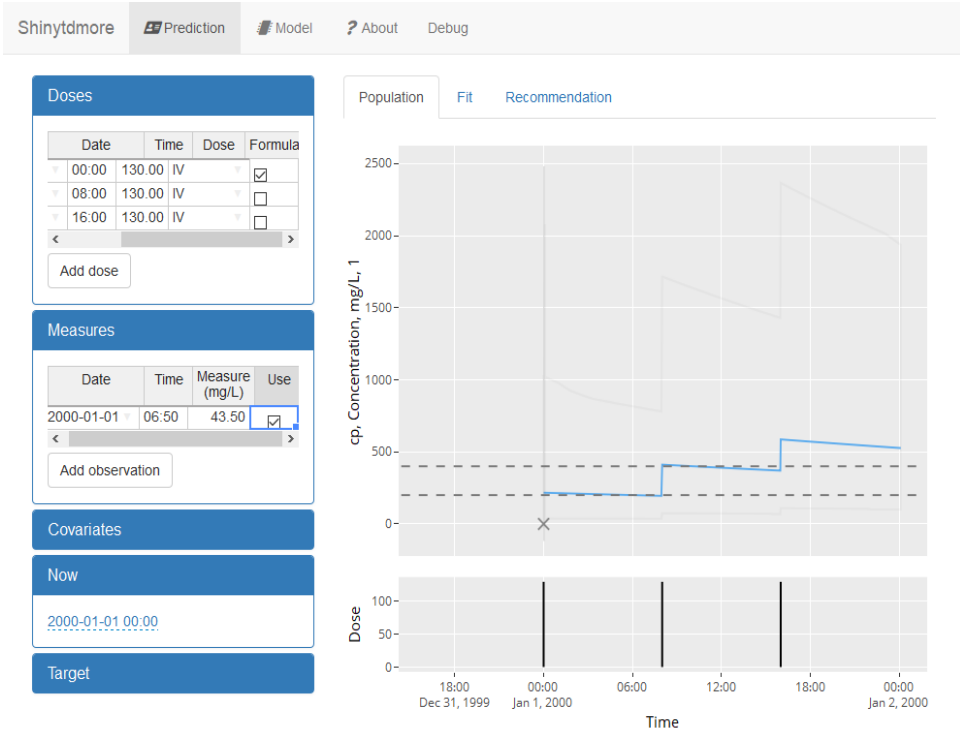


***Figure 19**: Shinytdmore standard interface showing (on the left) input tables for doses, measures, covariates, current time and target, and (on the right) typical value plot and dosing history.*

The `shinytdmore` R package was designed as self-contained user interface (UI) elements, a data model, and a reference Shiny application. The data model contains the treatment regimen, patient covariates, observed values, current time and target information. Time is represented as absolute calendar times (datetime), rather than numeric time since first dose.

All user interface elements are implemented as Shiny modules. The `handsontable` for editable tables was amended with editable datetime columns. A horizontal border was added as a visual element to distinguish past events from future events. This was further refined into a dosing regimen table, observations table, and covariates table. These tables are automatically generated based on the model metadata.

These elements were then assembled into larger functional blocks. The main `predictionTab` contains several collapsible tables on the left side for data input: dosing regimen, covariates, measures, and target. On the right, three views are available: population prediction, individual prediction, and dose recommendation. The population prediction plot allows a physician to visually inspect the dosing regimen and measured concentrations. Any typing errors, e.g. `150`mg instead of `15`mg, should be immediately apparent. The second view allows to see the individual fit. The third view shows the recommended dosing regimen and predicted concentrations.

This is complemented by the `reportTab`, allowing the user to download a report of all input data, model, and resulting dosing recommendation. The `aboutTab` displays all relevant version information, and the `modelTab` allows the physician to try alternate models.

For plots, the `plotly` interactive library is used. This allows to update plotted curves through animations, improving user experience. Plots also feature tooltips, zooming, panning and dynamic axis labels.

In a standard `shiny` application, any change in input values result in a recalculation of all output values, which blocks the user interface. Ideally, this recalculation is delayed until the user is done editing. This so-called *debouncing*[63] was implemented for all costly calculations, such as the estimation of an individual fit or dosing recommendation. Furthermore, additional logic was added to reduce calculation further: changing future doses does not require recalculation of the individual parameters.

For maintainability, custom javascript was kept to a minimum. This ensures the application can be kept up-to-date with new versions of shiny, handsontable or plotly with minimal effort. Package users or maintainers only require knowledge of R and Shiny, not of HTML or Javascript.

5.6. Stable and robust software

First, software should be maintainable. `tdmore` has been designed as modular software adhering to current software design patterns. Each function has a single responsibility. The software is well-documented: out of 5289 lines of code, only 2269 (43%) are effective code. 1430 lines (27%) are

[63] https://shiny.rstudio.com/reference/shiny/1.0.4/debounce.html

documentation, with the remainder being structural lines (whitespace or delimiters).

Documentation is further extended through the use of vignettes: examples of how to use the software. All code in a vignette is executed and results are included automatically in the documentation, either as plots or as screenshots of the user interface. This documentation is then made available on a public website[64].

In medical device software, stability is essential. At minimum, software should perform as intended. Automated testing was used to ensure this: a set of calculations is performed automatically and compared to a previously validated *reference* set. For 2269 effective lines of source code in `tdmore`, 2550 lines of tests were written, totaling 303 tests. These tests cover 85.46% of all code. Numerical imprecision of structural equation solving or numerical optimization was taken into account to reduce the number of false positives. Plotting functionality was also tested using the `vdiffr` package, which standardizes plot generation independent of the operating system, available fonts or screen size. This allows comparison of plots generated on a Linux server with the reference set generated on a Windows(TM) desktop machine. Interactive interfaces such as `shinytdmore` elements were tested using the automated web browser `phantomjs`. The application behavior and output was tested automatically by verifying either full HTML output, or specific interface elements. Formal computer systems validation (CSV), as described by US FDA, comprises installation qualification (IQ), operational qualification (OQ) and performance qualification (PQ). For the latter, execution time of tdmore calculations is compared to the execution time of a reference benchmark (execution of Escoufier's method on a 60x60 matrix), with a pre-defined allowed ratio.

This automated testing suite is executed automatically at every change in the source code, by a continuous integration server, using all latest versions of packages. Therefore, issues caused by an incompatibility between a new R version and `tdmore` are also detected automatically. In case of failure, the developers are notified through e-mail. As the `tdmore` framework may also be used in medical device software, a `stable` version of the software is also maintained. In this software version, a fixed version of dependent packages

[64] https://tdmore-dev.github.io/tdmore/dev/

and R base system is used through the `renv` system. This `stable` version is also tested automatically at every source code change.

For easy deployment, a Docker[65] image is also created automatically, tested, and published on a cloud repository. This allows fast, automated deployment of any tdmore-based application on cloud servers.

[65] Docker is a software platform that simplifies the process of building, running, managing and distributing applications. It does this by virtualizing the operating system of the computer on which it is installed and running. – Docker documentation

5.7. Tacrolimus application

5.7.1. Connection to EHR

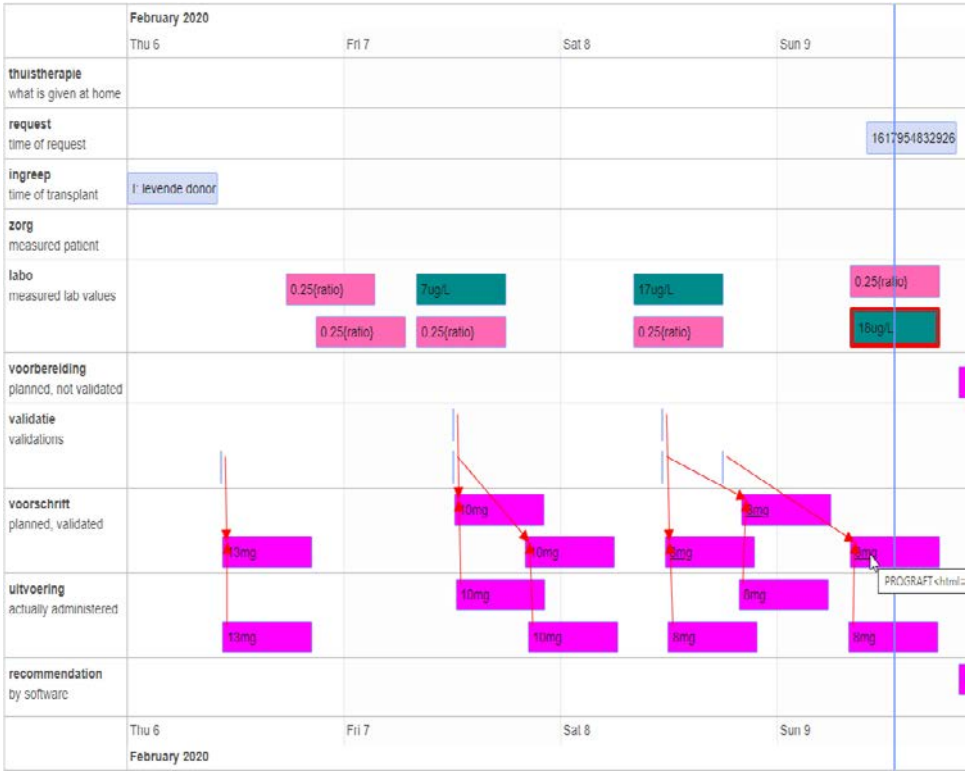


***Figure 20**: The timeline shows all data imported from the EHR in a compact overview*

To allow import of clinical data, a web Application Programmers Interface (API) was built using the `plumber` package. This allows the electronic health record (EHR) system to directly send data to the precision dosing application. The application parses this data and transforms it to a dataset fit for estimation and dose adaptation. This data consists of patient information (year of birth, sex), transplant information (transplant type, transplant date), current request (time and id), all known patient bodyweights, blood values (hematocrit and tacrolimus levels), and planned, validated and executed doses. The input is transformed into pharmacometric data compatible with `tdmore`, precision dosing is performed, and resulting dose recommendations

are returned to the EHR system. All input and output is stored in a database to allow future consultation by physicians through the web application.

In computer science, the “garbage in garbage out” principle states that flawed input data produces nonsense output. To distinguish between computation errors and data input errors in the case of erroneous dosing advice, a graphical visualization was made for all input data. This complemented the existing `shinytdmore` elements that allow introspection into the dosing advice.

5.7.2. Clinical workflow

To minimize this projects impact on the hospital, we aimed to fit within the current clinical workflow. For de novo kidney transplant recipients, blood is sampled in the morning before first administration, after which the first dose of tacrolimus is administered. The tacrolimus concentration in the blood sample is determined at the central lab. These results are made available in the EHR around noon. During rounds in the afternoon, the dose for each patient is adapted according to tacrolimus whole blood levels and (possibly) other factors.

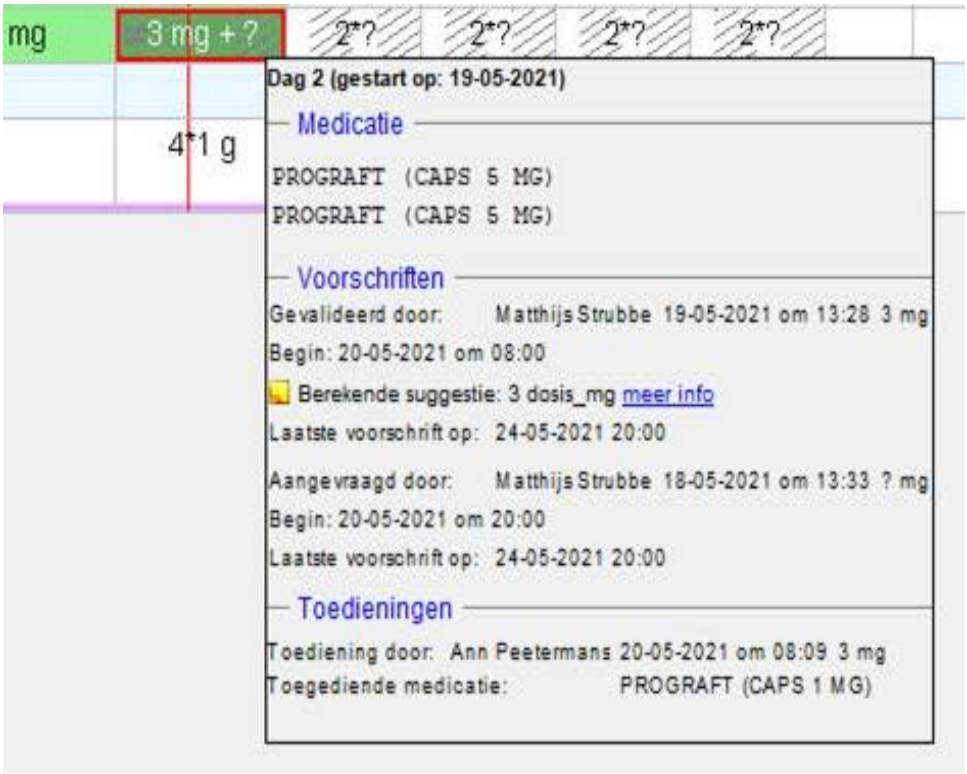


***Figure 21**: Screenshot of EHR interface with post-it showing 'Berekende suggestie: 3mg'. A hyperlink to the MIPD web application is displayed, allowing doctors to request more detailed information.*

To integrate MIPD into this workflow, we decided to present dosing advice as a "post-it" in the EHR system. This existing functionality required no adaptation to the EHR user interface. Only the EHR backend was adapted to allow the creation of post-its by the MIPD system.

A smooth automation was a fiendishly difficult task to achieve. The process is shown below in pseudo-code. Either a dosing advice (possibly with warnings) is returned to the EHR, or an error explaining why dosing advice cannot be provided. As a side-effect, e-mails are sent to the clinical team informing them of any issues. To limit the risk of notification overload, only relevant errors result in an email, e.g. a patient in the control arm will not generate any dosing advice, but this does not concern the clinical team.

```
In = receive()
json_validate(In)
```

```
if(In.study != "1234"): stop("Not in study")
storeInDatabase(In)

if(In.arm == "Control"): stop("Control arm")
if(In.arm not "Control" or "Intervention"): stop("Invalid ar
m")
Events = toEventList(In)
if(Events.trigger != "tacSample") stop("Not TAC trigger")
fixEvents(Events)
checkEvents(Events)

tdmoreIn = transformToTdmore(Events)
fit = estimate(tdmoreIn)
recommendation = findDoses(fit)

adaptInput(In, recommendation)

finally:
  if any error or warning is relevant:
    emailClinicalTeam()
```

The input is first validated using a JSON Validation schema[66]. We verify the patient is enrolled in the study, as no data may be stored from patients without informed consent. The study arm is verified, as this is a free-text field; using the wrong word (only "intervention" or "control" are allowed) results in an email to the clinical team. If the patient is in the intervention arm, the data input is transformed to a list of events `(Time, EventInfo)`. To limit the number of updates, dosing advice is only provided upon receiving a tacrolimus blood sample value, not when e.g. a new bodyweight is measured. This list is then further adapted based on the following business rules:

- Blood samples are registered in the EHR at their intended collection time (08:00 AM), not their actual collection time. In reality, blood is always sampled before administration of tacrolimus. If the EHR states tacrolimus was administered at e.g. 07:30, the blood sample collection time is then also adapted to 07:20 (10 minutes before administration).

[66] https://json-schema.org/

- Any blood sample times in the future are moved to the current time.
- Drug administrations other than tacrolimus are discarded.
- A previous prescription of "0mg" is considered administered, even if no administration was recorded.
- Any duplicate lab values or administrations are discarded
- The MIPD tool assumes prescriptions for the following day can be adapted. Prescriptions in the next 3 hours are left untouched. Prescriptions validated after the last lab measurement are left untouched; they are assumed "overridden" by the physician.

This list of events is then checked for any errors or inconsistencies. This step may result in a critical error -no dose advice is given- when giving automated dosing advice, but is de-escalated to a regular warning when performing dosing advice via the interactive web application.

- The transplant date should be known, and less than 14 days ago.
- There are no measurements/transplants/drug administrations in the future
- All past prescriptions were administered.
- Administrations without a corresponding prescription result in a warning.
- All tacrolimus administrations and prescriptions are the hard caps prograft or advagraf formulation. No IV or oral suspension is allowed.
- All lab values (hematocrit, tacrolimus whole blood level) have the right units.
- If tacrolimus was measured, a tacrolimus administration should precede it.
- Past prescriptions are compared to past dose recommendations. Any dose recommendation not followed results in a warning for the clinical team.

The above list was carefully crafted during 6 months of dry run. During that period, the software received information from the EHR and returned dosing advice. The dosing advice was never recorded in the EHR and never shown to physicians, but was analyzed by the study team for any inconsistencies. Thanks to this dry run period, most data errors could be detected, debugged and mitigated.

Now that the input data is correct, it is transformed to `tdmore` format. An individual fit is estimated, and this is used to adapt the treatment regimen. Finally, this recommendation is transformed back to the EHR format and returned as a result. It is amended with a unique hyperlink allowing access to the interactive web application. As a final step, any warnings or errors are sent to the clinical team via e-mail.

On average, about 50 requests were sent by the EHR to the MIPD application. Each request is stored in the database. A specific website was created that displays an overview of all requests to the clinical team. Personal information (transplant date, patient year of birth, patient sex) is shown to allow easy tracking of a specific patient across all requests.

An additional advantage of close integration with the EHR is in user authentication and authorization. The API endpoint can be restricted with a single pre-shared key included in the request header, and access to the management interface is only required for the study team. Access to the web application by physicians is instead managed by the EHR. The EHR manages user authentication, and whether a user is authorized to see the file of a specific patient. As long as the user can access the “More info” hyperlink in the EHR interface, the user is authorized to access the MIPD application for that patient.

5.7.3. Continuous deployment in the cloud

The tacrolimus application was deployed to the Google Cloud Platform. To support this, a webhook to a private git repository was set up. This triggers a Docker build each time the source code is modified, or when the supporting framework (shinytdmore, tdmore) is modified. The build process automatically runs smoke tests and updates a Kubernetes cluster definition to migrate to the new image. Upgrades result in 0 downtime, as the new version is started alongside the old version. Requests are routed to the new version, and the old version is stopped when all currently calculating dose recommendations have finished. To the best of our knowledge, this is the first

time continuous deployment has been documented for software where medical device regulations apply.

5.8. Legal challenges

5.8.1. Open-sourcing

Originally, the project aimed to publish an open-source framework, enabling other researchers to easily evaluate MIPD and build precision dosing software. As an added requirement, licensing of any precision dosing software built upon `tdmore` should remain flexible, and even allow for commercial distribution.

The intellectual property of `tdmore` is fully owned by the University of Leuven. This allows the university to distribute the software under any license it deems fit. The code for simulation was distributed under the `Afero GPLv3` license, a copy-left license. Any developer distributing software including `tdmore` is therefore obliged to allow users to download the source code as well, even if it only allows access to a web application built using `tdmore`. This answers a growing concern about closed-source medical devices, allowing public review and accountability of software with a potential far-reaching (deadly) impact. As the sole owner, the university also has the option to dual-license the software to potential commercially interested parties in a more permissive license.

A subset of the code was published on Github. Any user can fork the code and adapt their own version. To merge these improvements back into the official `tdmore` version, developers are required to transfer copyright of these improvements to the University of Leuven. This allows incorporation of these changes into any future closed-source dual-licensed versions.

To ensure the software can indeed be dual-licensed, a review of all dependent packages and their license terms was performed. All packages with MIT or BSD-style licenses can be included without restriction, as long as appropriate credit is given. For the `plotly` package, an older version was used whose license still allowed commercial use. However, many other dependent packages use the GNU Public License v2/v3, including base R itself. In principle, all code that *links* to GPL-covered software should also be

published under a GPL-compatible license (i.e. as open-source). However, this is the subject of much debate.[67]

> Whether or not they classify as one or two programs, or wether or not your package is a modified work of the dependencies is very much subject to interpretation and of the level of interactions between the packages. - Colin Fay, 2019

While the R Core Team is of the opinion that merely *using* R packages does not constitute "one program", and have publicly stated[68] they will not legally pursue any closed-sourced aggregate work, this is not a definitive stop to legal liability. RStudio has undertaken efforts to move many of their packages to MIT licenses instead, enabling commercial use without legal risk. In practice, `tdmore` heavily depends only on packages by RStudio, and only sporadically on other packages. No dependent packages are used that prohibit commercial use.

5.8.2. Patents

The distribution of `tdmore` under the Afero GPLv3 also requires shielding users from credible software patent threats. The methods used in `tdmore` may be construed to infringe upon the patent "System and method for providing patient-specific dosing as a function of mathematical models updated to account for an observed patient response", granted to Diane R Mould and Bayesient LLC by the US Patent Office in 2014. However, the validity of this patent is questionable. The US patent office grants claims without a reasonable search for prior art, and relies on competitors to challenge patents. In contrast, the European Patent Office performs a diligent search for prior art before granting a patent. The relevant patent application #EP2904388[69] was submitted in April '14, and was denied in April '15 as the claims lacked novelty. The claims were amended one week later. In May

---

[67] See https://thinkr-open.github.io/licensing-r/rlicense.html (consulted 19-NOV-21) for more details.

[68] https://cran.r-project.org/doc/FAQ/R-FAQ.html#Can-I-use-R-for-commercial-purposes_003f, consulted 19-NOV-21

[69] https://register.epo.org/application?number=EP13843939&lng=en&tab=doclist, consulted 19-NOV-2021

2016, European search opinion reconfirmed the lack of novelty. Over the course of 5 years since, the patent claim was amended an additional 5 times (December '16, October '17, September '19, November '19, and August '20). It appears the applicant aims to keep the claim open as long as possible, with no substantial progress to addressing the lack of novelty. We concluded that this patent is not a credible threat, and therefore a (user-transferable) patent license is not required to distribute the software.

5.8.3. Medical device legislation

`tdmore` is a framework for building precision dosing software. By providing individual predictions, this software is considered a Software Medical Device under both FDA and EMA guidances. For distribution in Europe, a medical device or medical device parts should have a CE approval. The requirements for CE validation are fundamentally incompatible with non-profit free-of-charge distribution. Among others, a liability structure, software helpdesk and safety follow-up is required. This requires continuous funding.

Could we simply distribute the software as-is, without warranty of any kind? In the US, this is indeed possible. However, Belgian and European law stipulate that a distributor cannot fully disclaim liability. The GPL clause disclaims all liability "to the extent permitted by law", and the EU specifically created the EU Public License to make this liability more precise:

> Article 8 - Disclaimer of liability: Except in the cases of willful misconduct or damages directly caused to natural persons, the Licensor will in no event be liable [...] - EUPL v1.1

Purposefully distributing a medical device without CE approval may indeed be regarded as willful misconduct, and liability directly caused to natural persons cannot be excluded. In the United States, software can be distributed without warranty of any kind. Consider the Open Artificial Pancreas System, an open-source software that measures blood sugar levels and drives an insulin pump. As users implement this system at their own risk, developers/distributors of the code have little to fear. There is much resistance to these systems however, with 91% of health care providers not

supporting the use of Do It Yourself systems for the administration of insulin in Type 1 diabetes.[70]

> Their (FDA) responsibility is to regulate products on the commercial market and help safeguard the public. OpenAPS is NOT a commercial product and is not sold or distributed in anyway. Individuals who build an OpenAPS are essentially doing an (n=1) experiment, which they have a right to do to/by themselves. That is not a regulated activity by the FDA. – OpenAPS.org public website

The lack of possibility to distribute software as-is in Europe is an insurmountable hurdle. When distributing open-source software, the European distributor is exposed to legal liability when this software is used for precision dosing by individuals. To accomplish our goals maximally within the confines of the current regulations, all functionality for individual fitting, prediction, or dose recommendation was made inaccessible. All documentation discussing this functionality was also removed. The software can only be used to perform research on MIPD performance, be it through retrospective datasets or fully virtual populations. Extensive documentation describing how to perform individual prediction and dose recommendation was permanently deleted from the source code repository. Of course, the code for individual predictions remains an integral and required part of the software, but accessing it is a conscious decision on part of the user. Both `shinytdmore` and the `tacrolimus` reference software tool could not be made available publicly.

5.8.4. General Data Protection Regulation (GDPR)

The implemented tacrolimus software as used in the clinical study was considered a medical device under evaluation, and therefore did not require a CE label. However, GDPR and data privacy regulations still applied to this clinical study. Usually, this is resolved by anonymizing all personally identifiable information. Patient data cannot be anonymized in this application though; the transplant date is a required covariate for the popPK model, as shown by in silico studies. Fortunately, the application could be hosted in the hospital IT infrastructure on Google Cloud Platform. The data never left the auspices of the hospital, thereby resolving the problem from a GDPR standpoint. Worryingly, this issue is generally ignored by commercial

[70] Kesavadev et al., "The Do-It-Yourself Artificial Pancreas"; Dickson et al., "#WeAreNotWaiting DIY Artificial Pancreas Systems and Challenges for the Law".

vendors of MIPD software, who ask physicians to input patient data on remote servers.

5.9. Conclusion

`tdmore` is a software package that supports research, development and implementation of MIPD. It is a flexible tool for research, yet is also robust and well-tested to the quality required for medical device software. We have demonstrated a smooth transition from research to implementation for tacrolimus dosing in de novo kidney transplant recipients. Although the software is not fully available due to legal challenges, we hope this cost-effective roadmap can be applied to other compounds in the future.

Chapter 6 Predicting infliximab MIPD for ulcerative colitis patients

This chapter is based on:

- Erwin Dreesen, Ruben Faelens, Gert Van Assche, Marc Ferrante, Séverine Vermeire, Ann Gils, and Thomas Bouillon. Optimising infliximab induction dosing for patients with ulcerative colitis. (2019) *British journal of clinical pharmacology*
- Faelens, Ruben, Ruben Faelens, Thomas Bouillon, Erwin Dreesen, Gert Van Assche, Marc Ferrante, Severine Vermeire, Ann Gils (2018), "Benefits of TDM in clinical management: a test case in infliximab for ulcerative colitis." *presented at PAGE 2018 conference as a scientific poster*
- Faelens Ruben, Zhigang Wang, Thomas Bouillon, Paul Declerck, Marc Ferrante, Séverine Vermeire, and Erwin Dreesen (2021). Model-informed precision dosing during infliximab induction therapy reduces variability in exposure and endoscopic improvement between patients. *Pharmaceutics, special issue "Model-Informed Precision Dosing"*., October 2021.

It plays an important part in the overarching thesis for three reasons. First, this is a population simulation. It shows how a virtual population can be used to compare outcomes for dosing strategies. Second, the model used is a PK/PD model. While model-informed precision dosing largely focuses on *probability of PK target attainment*, a PK/PD model allows us to directly evaluate relevant clinical outcomes instead. Finally, this study is important because it is a *negative result*. Through simulation, we showed the candidate dosing strategy would not improve clinical outcomes, thereby avoiding an expensive clinical study.

6.1. Abstract

Model-informed precision dosing (MIPD) may be a solution to therapeutic failure of infliximab for patients with ulcerative colitis (UC), as underexposure could be avoided, and probability of endoscopic improvement (pEI; Mayo endoscopic subscore ≤ 1) could be optimized. To investigate *in silico* whether this claim has merit, four induction dosing regimens were simulated: 5 mg/kg (label dosing), 10 mg/kg, covariate-based MIPD (fat-free mass, corticosteroid use, and presence of extensive colitis at baseline) and concentration-based MIPD (based on the trough concentration at day 14). Covariate- and concentration-based MIPD were chosen to target the same median area under the infliximab concentration-time curve up to endoscopy at day 84 (AUC_{d84}) as was predicted from 10 mg/kg dosing. Dosing at 5 mg/kg resulted in a mean$\pm$standard deviation pEI of 61.2% $\pm$ 55.7%. Increasing the dose to 10 mg/kg was predicted to improve pEI to 68.6% $\pm$ 65.1%. Covariate-based MIPD reduced variability in exposure and pEI (68.7% $\pm$ 65.1%). Concentration-based MIPD decreased variability further (69.3% $\pm$ 66%) but did so at an increased average dose of 2298mg per patient, as compared to 2181mg for 10 mg/kg dosing. Mean pEI remained unchanged between 10 mg/kg dosing and MIPD since the same median AUC_{d84} was targeted. In conclusion, quantitative simulations predict MIPD will reduce variability in exposure and pEI between patients with UC during infliximab induction therapy.

6.2. Introduction

Infliximab is a monoclonal antibody that binds and neutralizes the functional activity of tumor necrosis factor-alpha (TNFα). Based on the results of the landmark Active Ulcerative Colitis Trials (ACT) 1 and 2, infliximab was approved for inducing and maintaining remission in patients with moderate-to-severe ulcerative colitis (UC).[71] In these studies, endoscopic improvement (defined as Mayo endoscopic subscore ≤ 1) was achieved in about 60% of patients after administration of three infliximab infusions (5 mg/kg body weight, at weeks 0, 2, and 6; endoscopy at week 8). In post-marketing studies, endoscopic improvement rates were lower (e.g. 47% in Brandse et

[71] Rutgeerts et al., "Infliximab for Induction and Maintenance Therapy for Ulcerative Colitis".

al.[72]), making unpredictable outcomes of infliximab induction therapy a challenge[73].

Dose finding in ACT 1 and 2 failed to show consistent benefit of 10 mg/kg dosing over 5 mg/kg dosing.[74] However, higher infliximab serum concentrations during induction therapy were found to correlate with short-term endoscopic improvement, as well as long-term relapse-free, and colectomy-free survival.[75] To date, the infliximab exposure-response relationship in patients with UC has been well-established.[76] Consequently, it has been hypothesized that targeting infliximab to a predefined "optimal" exposure has the potential to improve the response rate and identify primary non-responders (defined as non-response despite optimal infliximab

---

[72] Brandse et al., "Pharmacokinetic Features and Presence of Antidrug Antibodies Associate With Response to Infliximab Induction Therapy in Patients With Moderate to Severe Ulcerative Colitis".

[73] Brandse et al.; Balzola et al., "Trough serum infliximab: A predictive factor of clinical outcome for infliximab treatment in acute ulcerative colitis: Commentary"; Papamichael et al., "Infliximab Concentration Thresholds During Induction Therapy Are Associated With Short-term Mucosal Healing in Patients With Ulcerative Colitis"; Farkas et al., "Efficacy of infliximab biosimilar CT-P13 induction therapy on mucosal healing in ulcerative colitis".

[74] Rutgeerts et al., "Infliximab for Induction and Maintenance Therapy for Ulcerative Colitis".

[75] Adedokun et al., "Association Between Serum Concentration of Infliximab and Efficacy in Adult Patients With Ulcerative Colitis".

[76] Papamichael et al., "Infliximab Concentration Thresholds During Induction Therapy Are Associated With Short-term Mucosal Healing in Patients With Ulcerative Colitis"; Arias et al., "A Panel to Predict Long-Term Outcome of Infliximab Therapy for Patients With Ulcerative Colitis"; Kobayashi et al., "First trough level of infliximab at week 2 predicts future outcomes of induction therapy in ulcerative colitis—results from a multicenter prospective randomized controlled trial and its post hoc analysis"; Vande Casteele et al., "Infliximab Exposure-Response Relationship and Thresholds Associated With Endoscopic Healing in Patients With Ulcerative Colitis".

exposure)[77]. To date, most therapeutic drug monitoring (TDM) studies of infliximab focus on maintenance therapy, whereas induction therapy is relatively unexplored. Moreover, the utility of TDM of infliximab in patients with UC remains controversial because of poor evidence from prospective TDM studies.[78] One potential reason for the weak evidence can be the use of inefficient TDM algorithms (analogous flowcharts and decision trees) in these TDM studies[79]. Therefore, model-informed precision dosing (MIPD), a more efficient and precise dose optimization strategy as compared to analogous TDM, has been suggested as a way out of the dilemma.[80]

MIPD can be implemented through either *a priori* or *a posteriori* dose optimization, both utilizing a population pharmacokinetic (popPK) model that serves as a prior. *A priori* dose optimization is done by involving patient's covariates/characteristics that explain between- and within-subject variability, while *a posteriori* dose optimization (Bayesian forecasting) is

---

[77] Brandse et al., "Pharmacokinetic Features and Presence of Antidrug Antibodies Associate With Response to Infliximab Induction Therapy in Patients With Moderate to Severe Ulcerative Colitis"; Vande Casteele et al., "American Gastroenterological Association Institute Technical Review on the Role of Therapeutic Drug Monitoring in the Management of Inflammatory Bowel Diseases"; Magro et al., "Third European Evidence-Based Consensus on Diagnosis and Management of Ulcerative Colitis. Part 1".

[78] Vande Casteele et al., "Trough Concentrations of Infliximab Guide Dosing for Patients With Inflammatory Bowel Disease"; Vande Casteele et al., "American Gastroenterological Association Institute Technical Review on the Role of Therapeutic Drug Monitoring in the Management of Inflammatory Bowel Diseases"; Mitrev et al., "Review article: consensus statements on therapeutic drug monitoring of anti-tumour necrosis factor therapy in inflammatory bowel diseases"; Syversen et al., "Effect of Therapeutic Drug Monitoring Vs Standard Therapy During Infliximab Induction on Disease Remission in Patients With Chronic Immune-Mediated Inflammatory Diseases".

[79] Wang and Dreesen, "Therapeutic Drug Monitoring of Anti-Tumor Necrosis Factor Agents".

[80] Keizer et al., "Model-Informed Precision Dosing at the Bedside"; Wang and Dreesen, "Therapeutic Drug Monitoring of Anti-Tumor Necrosis Factor Agents".

based on previous infliximab serum concentration measurements.[81] Through these two approaches, the MIPD software tool can recommend a dose that facilitates attainment of the therapeutic target exposure. Patient covariates such as C-reactive protein (CRP), serum albumin, antibodies to infliximab (ATI), body weight or fat-free mass, and fecal calprotectin have previously been identified in popPK modeling studies.[82]

In a previous popPK and exposure-response modeling analysis, we identified the relation between the cumulative area under the infliximab concentration-time curve up to endoscopy at day 84 (AUC_{d84}) and the probability of endoscopic improvement at day 84.[83] Based on these results, we suggested that increased exposures would result in better clinical outcomes. We further suggested that any increased drug consumption may be offset through the use of MIPD. In the present work, we investigated these claims further by performing population simulations of these different dosing scenarios and comparing exposures, probability of endoscopic improvement, and average drug consumption.

6.3. Materials and Methods

6.3.1. Population Pharmacokinetic and Exposure-Response Models

A previously published one-compartment popPK model with interindividual and interoccasion variability was used to simulate infliximab exposure.[84] This model was built on a total of 583 samples from 204 patients with UC,

---

[81] Vermeire et al., "How, When, and for Whom Should We Perform Therapeutic Drug Monitoring?"

[82] Vande Casteele et al., "Infliximab Exposure-Response Relationship and Thresholds Associated With Endoscopic Healing in Patients With Ulcerative Colitis"; Dreesen et al., "Optimising Infliximab Induction Dosing for Patients with Ulcerative Colitis," 2019.

[83] Dreesen et al., "Optimising Infliximab Induction Dosing for Patients with Ulcerative Colitis," 2019.

[84] Dreesen et al., "Optimising infliximab induction dosing for patients with ulcerative colitis," 2019.

and included C-reactive protein (CRP), serum albumin, and fat-free mass (FFM) as time-varying covariates, and Mayo endoscopic subscore, presence of extensive colitis, and corticosteroid use as baseline covariates.

Even though dose proportionality applies, when administering a higher dose of infliximab (cf. 6.3.3 *Dosing Scenarios*), a more positive disease evolution is expected, thereby influencing the time-course of CRP and serum albumin, both acute phase proteins, and possibly fat-free mass as well. Since the original dataset used for popPK model building did not include patients on higher infliximab doses (cf. 6.3.2 *Virtual Population*), and to avoid bias in the scenarios with higher dosing, we chose to re-estimate the model without these covariates. In theory, this should increase the unexplained interoccasion variability and residual error instead.

The logistic regression exposure-response model was adapted as well. The model was built on a subset of 159 patients and fitted the original data well.[85] However, this model predicted an ever-increasing probability of endoscopic improvement with increasing infliximab exposure. This could not be reconciled with the current line of thinking for infliximab treatment in UC, which assumes the existence of intrinsic non-responders.[86] The model was adapted to introduce maximum transition probabilities $E_{max,3\rightarrow 2}$ and $E_{max,2\rightarrow 1/0}$ for transitioning from a severe disease state (Mayo endoscopic subscore 3) to a moderate disease state (Mayo endoscopic subscore 2) and from a moderate disease state to endoscopic improvement (Mayo endoscopic subscore 1 or 0), respectively. Likelihood profiling was performed to identify the confidence bound for these parameters.[87] These E_{max} parameters were varied across a wide range of values and the associated AUC_{50}s (i.e., the infliximab exposures required to achieve half-maximal transition probabilities) were estimated, yielding a log-likelihood (LL) estimate for each parameter set. Estimates with $\Delta 2LL = 3.84$ showed the lower 95% confidence bound for the exposure-response model. These parameter estimates were then used for subsequent simulations.

---

[85] Dreesen et al.

[86] Ben-Horin, Kopylov, and Chowers, "Optimizing Anti-TNF Treatments in Inflammatory Bowel Disease".

[87] Sheiner, "Analysis of Pharmacokinetic Data Using Parametric Models. III. Hypothesis Tests and Confidence Intervals".

6.3.2. Virtual Population

To construct the virtual population for the dosing simulations, the original clinical dataset was used.[88] Only patients with a baseline Mayo endoscopic subscore of 2 or 3 were included, resulting in a source dataset of 194 patients. This dataset was expanded through Monte Carlo sampling of interindividual variability (200 samples per individual patient), yielding a total of 38,800 virtual patients. Baseline covariates were collected in a study conducted in accordance with the principles of good clinical practice and the Declaration of Helsinki. All patients provided written informed consent prior to participation in the Ethics Committee-approved IBD Biobank [B322201213950/S53684], whereby patients' characteristics and samples were collected prospectively on a series of predefined time points.

6.3.3. Dosing Scenarios

Four distinct dosing scenarios were evaluated. First, a standard dosing regimen of 5 mg/kg at days 0, 14, and 42 was applied to all virtual patients. Based on the exposure-response analysis of the original dataset, there was support for a higher dose.[89] Therefore, 10 mg/kg was evaluated as a second dosing scenario.

We aimed for covariate-based and concentration-based MIPD to result in the same mean predicted probability of endoscopic improvement as in the 10 mg/kg dosing scenario. Therefore, MIPD scenarios were designed to target the same median AUC_{d84} as was predicted from the 10 mg/kg dosing scenario. The third dosing scenario was purely based on the covariates (*a priori* MIPD). The popPK model was used to determine the covariate-based dose required to hit the exposure target associated with the predefined probability of endoscopic improvement.

Finally, Bayesian forecasting (*a posteriori* MIPD) was evaluated as a fourth dosing scenario. The sampled interindividual variability was used to simulate the trough concentration on day 14 resulting from the covariate-based first dose. Residual error was sampled and added to this concentration. This

[88] Arias et al., "A Panel to Predict Long-Term Outcome of Infliximab Therapy for Patients With Ulcerative Colitis".

[89] Dreesen et al., "Optimising infliximab induction dosing for patients with ulcerative colitis," 2019.

simulated concentration was subsequently used to perform an empirical Bayesian estimation of the patient's individual PK parameters. These individual estimates were then used to adapt the subsequent doses at days 14 and 42. Both doses were adapted to the same value, predicted to result in an AUC_{d84} exposure metric resulting in the target probability of endoscopic improvement.

6.3.4. Evaluation of Dosing Scenarios

The mean dose per patient and resulting exposures (AUC_{d84}) in each scenario were evaluated graphically as density plots. To quantify efficacy, the mean probability of endoscopic improvement was evaluated, as this reflects the expected fraction of patients attaining endoscopic improvement. Additionally, the mean overall dose per patient was evaluated. Finally, a robustness analysis was performed to determine whether our conclusions hold for other E_{max} parameter values.

6.3.5. Software

The adapted popPK and exposure-response models were estimated using NONMEM (version 7.4.3; Icon Development Solutions, Gaithersburg, Maryland, USA). Simulation of the dosing scenarios was performed using R (version 4.0.2; R Foundation for Statistical Computing, R Core Team, Vienna, Austria) with RxODE[90] and tdmore. The tdmore R package was developed at KU Leuven to perform simulation and evaluation of MIPD. It is available as open-source at *github.com/tdmore-dev/tdmore*. The NONMEM code and tdmore R code are provided in the Supplementary File.

6.4. Results

6.4.1. Population Pharmacokinetic and Exposure-Response models

The popPK model was adapted to include only covariates at baseline. As expected, the interindividual variability on the elimination rate constant and the proportional residual error increased (Table **9**). A visual predictive check of the updated popPK model is available in Figure **47**.

[90] Fidler et al., *RxODE*.

Likelihood profiling of the exposure-response model showed a wide range of probable $E_{max,3\rightarrow2}$-$E_{max,2\rightarrow1,0}$ pairs. In Figure **48**, the likelihood profile is shown for different parameter combinations. The $\Delta 2LL = 3.84$ contour line in red shows parameter combinations limits for $E_{max,3\rightarrow2}$-$E_{max,2\rightarrow1,0}$ of either 92.6%/100% or 100%/78.4%. Figure **22** shows the simulated PD model at parameter estimates with associated $\Delta 2LL = 3.84$. Based on this plot, $E_{max,2\rightarrow1,0} = 78.4\%$-$E_{max,3\rightarrow2} = 100\%$ was selected for further simulations, as the most "pessimistic" scenario. The remainder of possible parameter values were explored in the sensitivity analysis.

6.4.2. Dosing Simulations: Exposure and Efficacy

Simulation results are summarized in Table **2** and will be presented hereafter as median [95% prediction interval] for exposure (AUC_{d84}) and mean $\pm$standard deviation for probability of endoscopic improvement. As exposure and efficacy targets differ depending on the baseline endoscopic disease severity, results are reported for baseline Mayo endoscopic subscores of 2 (moderate disease severity; reported first) and 3 (high disease severity; reported second) separately. Exposures, mean doses and associated probabilities of endoscopic improvement are shown in Figure **23** and Figure **24**.

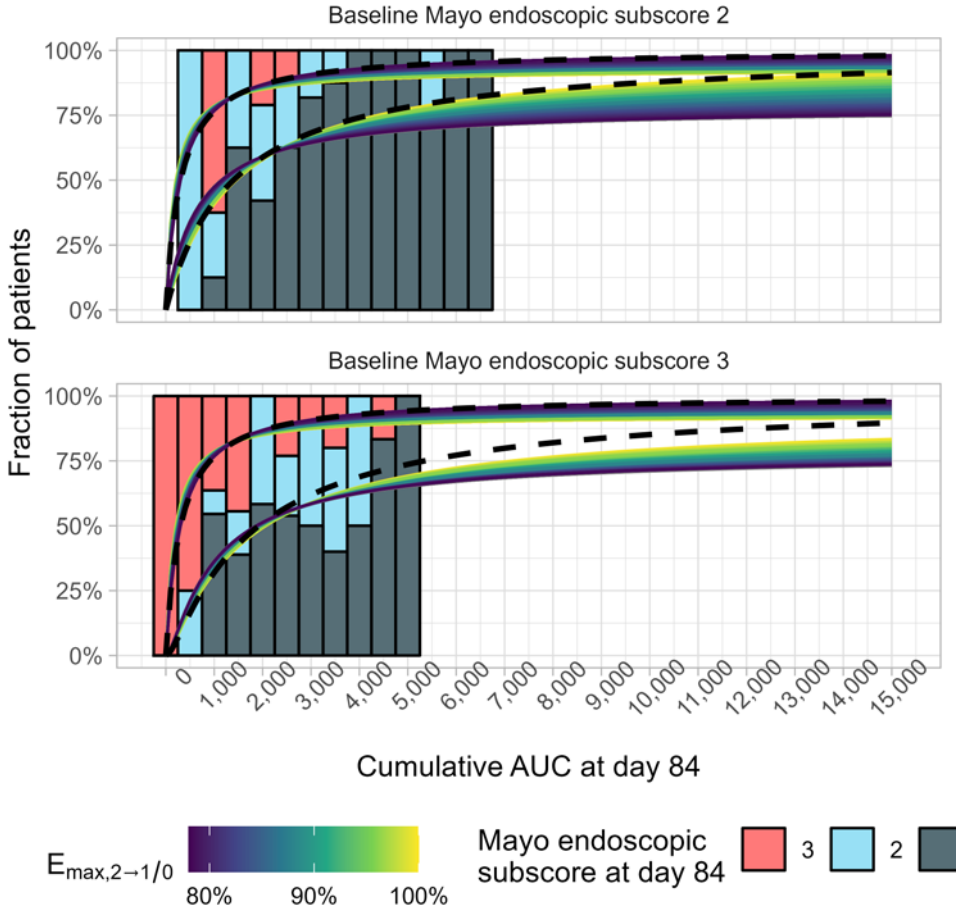


***Figure 22**: Exposure-response dataset (bars), binned per cumulative area under the curve (AUC) at day 84 and categorized according to Mayo endoscopic subscore at day 84, and corresponding simulated exposure-response models (lines representing the fraction of patients achieving a Mayo endoscopic subscore ≤ 1 [lower line] and ≤ 2 [upper line]). The original exposure-response model of Dreesen et al. is shown as a black dotted line. Colored lines represent models at $\Delta 2LL = 3.84$ with different $E_{max,2\rightarrow 1/0}$ values. All presented models fit the exposure-response dataset equally well (at $\alpha = 0.05$) but have different predictions outside the observed exposure range.*

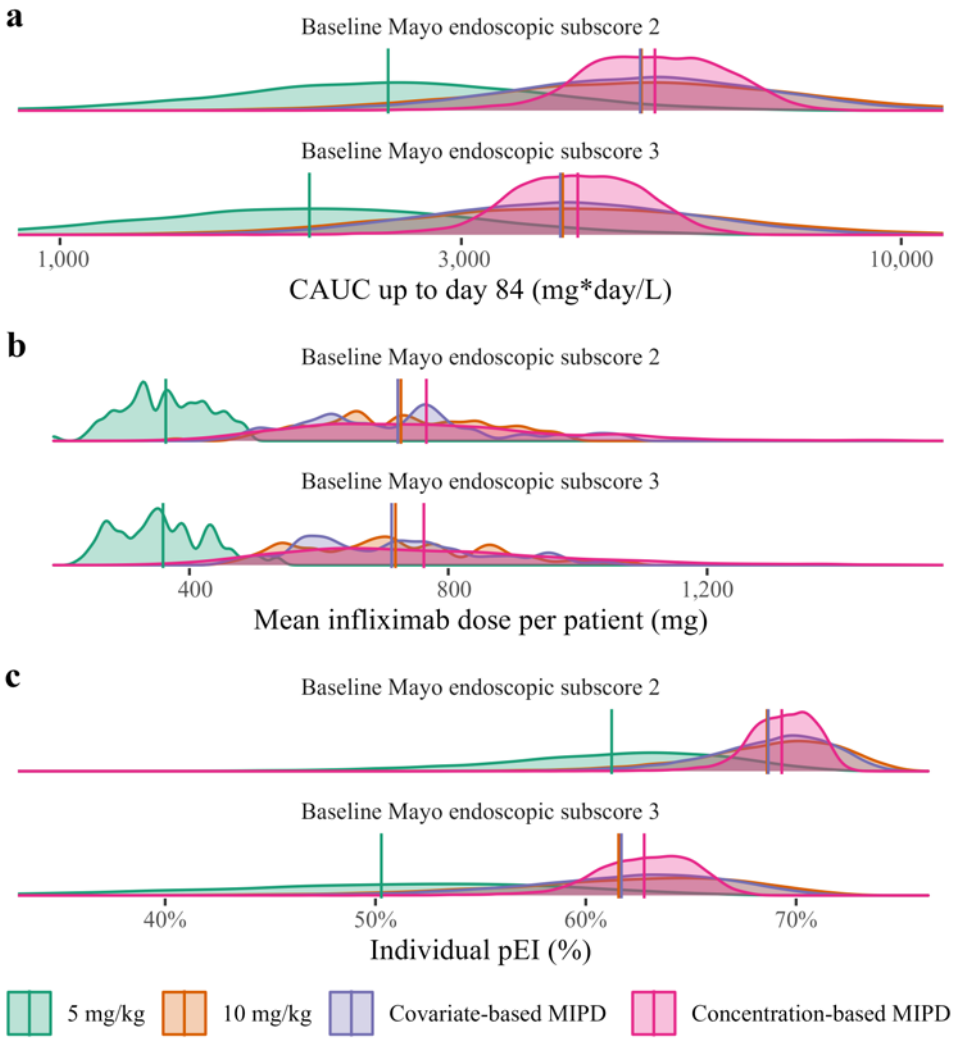


***Figure 23**: (a) Density plots of exposure in each of the four dosing scenarios, per baseline Mayo endoscopic subscore. Vertical lines show median exposure per scenario. (b) Density plots of the mean doses in each of the four dosing scenarios, per baseline Mayo endoscopic subscore. Vertical lines show overall mean dose per scenario. (c) Density plots of the individual probability of endoscopic improvement in each of the four dosing scenarios, per baseline Mayo endoscopic subscore. Vertical lines show overall mean pEI per scenario. CAUC, cumulative area under the curve; MIPD, model-informed precision dosing; pEI, probability of endoscopic improvement.*

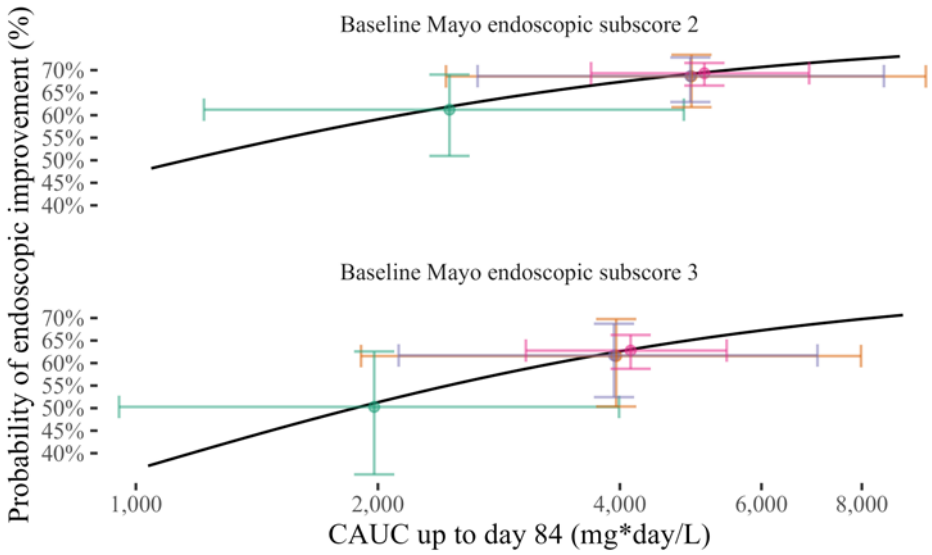


***Figure 24**: Predicted median exposure, mean probability of endoscopic improvement and 90% prediction interval in each scenario, per baseline Mayo endoscopic subscore. Black lines show model-predicted response. CAUC, cumulative area under the curve; MIPD, model-informed precision dosing.*

The 5 mg/kg dosing scenario resulted in an AUC_{d84} of 2455 [1215-4805] mg×day/L and 1979 [953-3990] mg×day/L, for baseline Mayo endoscopic subscore 2 and 3, respectively. This resulted in a predicted probability of endoscopic improvement of 61.2% ± 5.5% and 50.3% ± 8.4%. By increasing the dose to 10 mg/kg, exposure doubled to 4910 [2431-9609] mg×day/L and 3958 [1906-7981] mg×day/L. Probabilities of endoscopic improvement also increased to 68.6% ± 3.6% and 61.6% ± 6%.

Adapting the dose based on relevant covariates allowed more precise dosing, as between-population-variability can be taken into account. As can be seen in Figure **23**a, covariate-based MIPD resulted in the same median exposure as 10 mg/kg dosing, at a reduced variability (4895 [2661-8522] mg×day/L and 3933 [2123-7045] mg×day/L, for baseline Mayo endoscopic subscore 2 and 3, respectively). Dose adaptation based on the trough concentration measured at day 14 (Bayesian forecasting) further reduced this variability (5095 [3683-6879] mg×day/L and 4125 [3056-5431] mg×day/L). Probability of endoscopic improvement followed a similar pattern, with similar mean probabilities across 10 mg/kg dosing, covariate-based MIPD, and concentration-based MIPD.

***Table 2**: Summary of the simulation results*

		$AUC_{d84}(mg/L*day)$		$pEI(\%)$		$Cumulativedose(mg)$	
Baseline Mayo endoscopic subscore	Dosing scenario	median	[90%PI]	mean	±sd	mean	±sd
2	5 mg/kg	2,455	[1215-4805]	61.2	±5.51	1,090	±196
	10 mg/kg	4,910	[2431-9609]	68.6	±3.6	2,181	±393
	Covariate-based MIPD	4,895	[2661-8522]	68.7	±3.08	2,166	±443
	Concentration-based MIPD	5,095	[3683-6879]	69.3	±1.67	2,298	±613
3	5 mg/kg	1,979	[953-3990]	50.3	±8.36	1,078	±214
	10 mg/kg	3,958	[1906-7981]	61.6	±6.05	2,155	±428
	Covariate-based MIPD	3,933	[2123-7045]	61.7	±5.06	2,137	±417
	Concentration-based MIPD	4,125	[3056-5431]	62.8	±2.51	2,287	±643
Combined (2:3, 49%:51%)	5 mg/kg	2,210	[1049-4448]	55.7	±8.96	1,084	±205
	10 mg/kg	4,419	[2098-8895]	65.1	±6.11	2,168	±411
	Covariate-based MIPD	4,372	[2302-7940]	65.1	±5.46	2,151	±431
	Concentration-based MIPD	4,561	[3209-6516]	66.0	±3.91	2,293	±628

The systematically lower exposure at a baseline Mayo endoscopic subscore of 3 (severely active ulcerative colitis), as compared to a baseline Mayo endoscopic subscore of 2 (moderately active ulcerative colitis), may mechanistically be explained by a higher target load (target-mediated drug disposition) and protein-losing enteropathy (fecal drug loss). AUCd84, the area under the infliximab concentration-time curve from baseline up to endoscopy at day 84 (week 12); MIPD, model-informed precision dosing; pEI, probability of endoscopic improvement; PI, prediction interval; q, quantile; sd, standard deviation.

6.4.3. Dosing Simulations: Drug Consumption

Looking at the average infliximab dose used per patient (see also Figure **23**b), 5 mg/kg dosing resulted in 1090 mg per patient, and 10 mg/kg dosing doubled the dose usage to 2181 mg per patient. Covariate-based MIPD used an average of 2166 mg per patient. Concentration-based MIPD used at average 2298 mg per patient.

6.4.4. Sensitivity Analysis

The analysis presented above assumed an E_{max} plateau of 78% for the probability of transitioning from a Mayo endoscopic subscore of 2 (moderate disease severity) to a Mayo endoscopic subscore of 0 or 1 (endoscopic improvement). This plateau benefited MIPD, as overexposed patients were dose-reduced without significantly reducing probability of endoscopic improvement, while underexposed patients were dose-increased, thereby significantly increasing probability of endoscopic improvement.

Repeating our simulation study with higher values for $E_{max,2\rightarrow1,0}$ decreased this benefit, further favoring 10 mg/kg dosing, as is illustrated in Figure **22**. Other parameter combinations at $\Delta 2LL = 3.84$, as well as for the base model ($\Delta 2LL = 0$), consistently showed less favorable results for MIPD.

At $E_{max,3\rightarrow2}/E_{max,2\rightarrow1,0}$ of 92.6%/100%, the probability of endoscopic improvement for 10 mg/kg dosing was 77.1% [62.7%-86.5%] and 65% [48.1%-76.7%], for baseline Mayo endoscopic subscore 2 and 3, respectively, at an average drug consumption of 2181 mg per patient. Bayesian forecasting resulted in a probability of endoscopic improvement of 77.6% [71.7%-82.3%] and 65.8% [59.4%-70.9%], for baseline Mayo endoscopic subscore 2 and 3, at an average drug consumption of 2301 mg per patient.

6.5. Discussion

In this study, we compared four possible dosing scenarios for infliximab induction therapy in patients with UC: 5 mg/kg weight-based dosing (label dosing), and three dosing strategies with increased exposure: 10 mg/kg weight-based dosing, covariate-based MIPD, and concentration-based MIPD, all with unchanged timing of the infusions (day 0, 14, and 42). The 10 mg/kg dosing scenario was predicted to significantly improve endoscopic outcomes as compared to 5 mg/kg dosing. By design, MIPD (based on either covariates or the day 14 trough concentration) resulted in the same median exposure (AUC_{d84}) and, consequently, the same mean probability of endoscopic

improvement as observed in the 10 mg/kg dosing scenario. MIPD was predicted to successfully adapt individual patient doses, reducing the interindividual variability in infliximab exposure and with it the probability of endoscopic improvement, thereby providing "more equal" chances of endoscopic remission to all patients. Surprisingly, it did so at a higher average drug consumption per patient. Underexposed patients indeed received a relative dose increase, while overexposed patients received a relative dose decrease. However, this is a non-zero sum, as e.g., 10 mg×2+10 mg×0.5 > 10 mg+10 mg. Therefore, our simulation study showed improved outcomes under 10 mg/kg dosing as compared to 5 mg/kg dosing and showed additional benefit of MIPD over 10 mg/kg for reducing variability in exposure and efficacy between patients, however, at a higher direct drug cost. Performing MIPD may thus require a willingness to "pay for equality" amongst patients.[91] Consequently, we may consider shifting focus from outcome rates at the population level (the traditional industry perspective) to outcome chances at the individual patient level. It is at the individual patient level that MIPD may show value. Since the majority of patients attain the target under empirical dosing, it is important that future MIPD studies are restricted to vulnerable populations, such as patients with acute severe ulcerative colitis.[92]

The simulations described in this work were based on a previously established popPK model and exposure-response model of infliximab.[93] These models described the relation between the infliximab dose and exposure, and the AUC_{d84} and the probability of endoscopic improvement at day 84, respectively. The models were established on a dataset of 204 patients with moderate-to-severe UC. Since the majority of the infliximab doses in the original cohort were 5 mg/kg (approximately 90%), and only about 10% of the doses were 10 mg/kg, it should be noted that the exposure-response model was built on a relatively limited range of exposures. The exposures simulated in the present work exceed this range. However, this weakness was mitigated by a thorough analysis of exposure-response model

[91] Dreesen, "New Tools for Therapeutic Drug Monitoring".

[92] Battat et al., "Baseline Clearance of Infliximab Is Associated With Requirement for Colectomy in Patients With Acute Severe Ulcerative Colitis".

[93] Dreesen et al., "Optimising infliximab induction dosing for patients with ulcerative colitis," 2019.

parameter confidence intervals and likelihood profiling, and a sensitivity analysis. Our findings hold throughout the full range of probable model parameter values.

The exposure-response model assumes a causal effect between exposure and response. Previous clinical studies have indeed found a correlation between low trough concentrations and primary nonresponse to anti-TNFα therapy.[94] However, the causality assumed in our exposure-response model was never established in clinical studies. In light of this, time-varying disease-related covariates may instead be simulated in a joint model, avoiding potential underestimation of exposure at higher doses and reduced disease severity. Further research is needed to definitively establish whether non-response at low trough concentrations is due to mechanistic failure (pharmacodynamic [PD] failure) or underexposure (PK failure), as others have attempted to model this distinction.[95] Underexposure can be resolved through dose increase, while the mechanistic failure suggests switching to a different drug with another mechanism of action. A more fine-grained model of continuous endpoints may distinguish between PK and PD failure.

High exposure to infliximab may pose safety concerns. The 10 mg/kg dosing may result in very high exposures, which were predicted in the present study to be beneficial to patients. In reality, these highly exposed patients may present with adverse drug reactions such as infections, especially in the elderly, and MIPD may benefit these patients by reducing toxicity.[96]

[94] Ding, Hart, and De Cruz, "Systematic Review".

[95] Dreesen et al., "Modelling of the Relationship Between Infliximab Exposure, Faecal Calprotectin and Endoscopic Remission in Patients with Crohn's Disease"; Brekkan et al., "A Population Pharmacokinetic-Pharmacodynamic Model of Pegfilgrastim".

[96] Kantasiripitak et al., "The Effect of Aging on Infliximab Exposure and Response in Patients with Inflammatory Bowel Diseases"; Bejan-Angoulvant et al., "Brief Report"; Landemaine et al., "Cumulative Exposure to Infliximab, But Not Trough Concentrations, Correlates With Rate of Infection".

Our findings seemingly contradict the pivotal ACT 1 and 2 trials,[97] which showed no significant difference between 5 mg/kg and 10 mg/kg in endoscopic improvement rates on day 56 of therapy. Nevertheless, in the post-hoc PK-PD analysis of ACT 1 and 2, the exposure-response relationship has been established.[98] It would be worthwhile to repeat the presented modeling and simulation exercise including the data from these pivotal trials. Notwithstanding these results, clinical trials are currently underway evaluating an intensified induction regimen of 10 mg/kg.[99]

MIPD of infliximab has been implemented in clinical practice mainly in tertiary care centers, but even there, the confidence in MIPD is crumbling as the results of the landmark TAXIT, TAILORIX, and NOR-DRUM trials do not live up to expectations.[100] Our research showed that *in silico* simulations are a low-cost alternative to these clinical studies. Nevertheless, the translation of findings from a virtual trial into the real world may be challenged by noise due to for example sampling and measurement errors, rounding of doses and dosing intervals, etc.[101]

MIPD is classically used to improve probability of target attainment, with the target window defined by efficacy and toxicity. In this context, efficacy is

---

[97] Rutgeerts et al., "Infliximab for Induction and Maintenance Therapy for Ulcerative Colitis".

[98] Adedokun et al., "Association Between Serum Concentration of Infliximab and Efficacy in Adult Patients With Ulcerative Colitis".

[99] Austin Health and University of Melbourne, "Optimising Infliximab Induction Therapy for Acute Severe Ulcerative Colitis (PREDICT-UC)".

[100] Vande Casteele et al., "Trough Concentrations of Infliximab Guide Dosing for Patients With Inflammatory Bowel Disease"; Syversen et al., "Effect of Therapeutic Drug Monitoring Vs Standard Therapy During Infliximab Induction on Disease Remission in Patients With Chronic Immune-Mediated Inflammatory Diseases"; D'Haens et al., "Increasing Infliximab Dose Based on Symptoms, Biomarkers, and Serum Drug Concentrations Does Not Increase Clinical, Endoscopic, and Corticosteroid-Free Remission in Patients With Active Luminal Crohn's Disease".

[101] Alihodzic et al., "Impact of Inaccurate Documentation of Sampling and Infusion Time in Model-Informed Precision Dosing".

ever-increasing with higher exposures, while a dose increase is largely limited by cost rather than toxicity. It may be interesting to quantify the effect of infliximab as quality-adjusted life years (QALY) instead, allowing a direct comparison to increased cost and a straightforward optimization of QALY/cost.

In summary, we performed simulations to illustrate and predict the impact of three dosing strategies for increasing infliximab exposure during induction therapy as compared to 5 mg/kg weight-based label dosing, thereby improving the probability of endoscopic improvement. The use of 10 mg/kg dosing was indeed predicted to improve the probability of endoscopic improvement to 68.6% at an average drug consumption of 2181 mg per patient during induction therapy. Individualized dose adaptation could maintain the same mean probability of endoscopic improvement while reducing variability between individual patients. Although MIPD showed benefit for reducing variability in exposure and efficacy between patients, this comes at a higher direct drug cost as compared to 10 mg/kg weight-based dosing.

Chapter 7 Quantifying the impact of MIPD on endpoints: a test-case for tacrolimus

This chapter is based on:
Faelens Ruben, Nicolas Luyckx, Dirk Kuypers, Thomas Bouillon, and Pieter Annaert (2021). Predicting model-informed precision dosing: a test-case in tacrolimus dose adaptation for kidney transplant recipients. *CPT: Pharmacometrics and Systems Pharmacology*, submitted May 2021, accepted December 2021.

All description of clinical trial simulation were moved to Chapter 1, where these simulations are discussed in detail.

This work quantifies how MIPD improves PTA in tacrolimus. This demonstrates the key argument of this thesis work: investment into MIPD should only be performed after demonstrating the benefit in silico. This work further demonstrates our insights into the fine details of developing models for precision dosing. It shows how models can be evaluated and improved, how model selection should be performed for MIPD, and how the estimation method may be improved to incorporate parameter drift.

7.1. Abstract

Before investing resources into the development of a precision dosing (MIPD) tool for tacrolimus, the performance of the tool was evaluated in silico. A retrospective dataset of 315 de novo kidney transplant recipients was first used to identify a 1-compartment pharmacokinetic (PK) model with time-dependent clearance. MIPD performance was subsequently evaluated by calculating error to predict future concentrations, which is directly related to dosing precision and probability of target attainment (PTA).

Based on the identified model residual error, the theoretical upper limit was 45% PTA for a target of 13.5ng/mL and an acceptable range of 12 to 15 ng/mL. Using empirical Bayesian estimation, this limit was reached on day 5 post-transplant and beyond. By incorporating correlated within-patient variability when predicting future individual concentrations, PTA improved beyond the theoretical upper limit. This yielded a Bayesian feedback dosing algorithm accurately predicting future trough concentrations and adapting each dose to reach a target concentration.

Simulated concentration-time profiles were then used to quantify MIPD-based improvement on three endpoints: average PTA increased from 28% to 39%, median time to 3 concentrations in target decreased from 10 days to 7 days, and mean log-squared distance to target decreased from 0.080 to 0.055. A study of 200 patients was predicted to have sufficient power to demonstrate these nuanced PK endpoints reliably.

These simulations supported our decision to develop a precision dosing tool for tacrolimus and test it in a prospective trial.

7.2. Introduction

Optimal dosage of drugs with a narrow therapeutic index is an active area of research. These drugs display high pharmacokinetic (PK) or pharmacodynamic (PD) variability, exceeding the safe and effective variability[102] and therefore require individual dose adaptation. Covariate-based dose adaptation may be attempted first, as e.g. patient bodyweight can be easily measured. In case of large unexplained variability however, regular

[102] Holford and Buclin, "Safe and Effective Variability—A Criterion for Dose Individualization".

follow-up of the patient is needed to adapt the dose. In its simplest form, the drug is titrated until efficacy and safety is reached for the patient. Typically, drug concentration is used as a quantitative surrogate, aiming for an exposure previously established as having sufficient efficacy and acceptable safety.

One such drug requiring blood concentration monitoring is tacrolimus. At sufficiently high concentrations, tacrolimus acts as an immunosuppressor reducing the risk of graft rejection in solid organ transplant recipients.[103] High tacrolimus concentrations are strongly associated with a higher rate of drug-induced side effects however, including acute kidney injury.[104] The acceptable range depends on transplanted organ, immunological risk, adjunct immunosuppressive therapies, and may be further adapted to physician discretion.[105]

The required doses to achieve concentrations at a given target are highly variable between patients. Tacrolimus exhibits a high inter-individual PK variability,[106] with some studies reporting a 10-fold range in individual clearance.[107] To achieve safe and effective drug concentrations, regular follow-up of tacrolimus concentrations is therefore recommended.

However, translating an observed drug concentration into the required dose adaptation is not trivial. In theory, model-informed precision dosing (MIPD) can provide accurate dosing recommendations, yielding a high target

---

[103] Vincenti et al., "A Long-Term Comparison of Tacrolimus (Fk506) and Cyclosporine in Kidney Transplantation".

[104] Miano et al., "Early Tacrolimus Concentrations After Lung Transplant Are Predicted by Combined Clinical and Genetic Factors and Associated With Acute Kidney Injury".

[105] Brunet et al., "Therapeutic Drug Monitoring of Tacrolimus-Personalized Therapy"; Wallemacq et al., "Opportunities to Optimize Tacrolimus Therapy in Solid Organ Transplantation".

[106] Vanhove, Annaert, and Kuypers, "Clinical Determinants of Calcineurin Inhibitor Disposition".

[107] Brooks et al., "Population Pharmacokinetic Modelling and Bayesian Estimation of Tacrolimus Exposure".

concentration achievement. Many software packages offer dose adaptation of tacrolimus,[108] although only a single study has prospectively investigated whether MIPD improves this probability of target attainment (PTA).[109] Størset et al managed to demonstrate MIPD increases PTA, as compared to standard of care. Unfortunately, no improvement early post-transplant could be shown. The study may have been underpowered to show this, as only 80 patients were recruited. Enrollment was based on convenience, rather than a power calculation.

To evaluate the performance of MIPD, some *in silico* methods are available. Classical Goodness of Fit (GoF) metrics such as mean prediction error or root mean squared prediction error (RMSE%) show how well a model can fit existing data. Recently, prospective evaluation was proposed to evaluate these metrics on future concentrations, based on only the concentration samples collected up to that point.[110] Unfortunately, none of these metrics directly translate into the expected individual PK, PD or clinical outcomes when using MIPD.

Simulated data are preferred over GoF metrics for two reasons. Firstly, predicted individual data can be condensed to clinical benefit (e.g. avoidance of rejection), and this benefit is weighed against the implementation cost. Should pharmacogenetic information be included, how many additional blood samples are needed, and should MIPD even be implemented at all? Quantifying the potential clinical benefit is key in justifying investment into MIPD.

Second, simulated data help to design a prospective clinical trial comparing MIPD to standard of care. Without an *in silico* prediction of study endpoint, the only recourse is either to select the study sample size based on available

---

[108] Woillard et al., "Population Pharmacokinetics and Bayesian Estimators for Refined Dose Adjustment of a New Tacrolimus Formulation in Kidney and Liver Transplant Patients".

[109] Størset et al., "Improved Tacrolimus Target Concentration Achievement Using Computerized Dosing in Renal Transplant Recipients-a Prospective, Randomized Study".

[110] Lindbom, Pihlgren, and Jonsson, "PsN-Toolkit—A Collection of Computer Intensive Statistical Methods for Non-Linear Mixed Effect Modeling Using NONMEM".

resources (as large as we can), or on vague assumptions (we estimate X% improvement in PTA). Neither method is a solid way to design a clinical study, leading to inconclusive results. Even if there may be a theoretical and worthwhile benefit to MIPD, it may not be possible to demonstrate this effect in a reasonably funded clinical study. PTA predictions allow informed and realistic clinical study design.

This work shows how the impact of a Bayesian feedback MIPD tool optimizing PK outcomes can be predicted in tacrolimus dosing of renal transplant recipients the first 14 days post-transplant. These predictions are first used to evaluate whether the proposed MIPD achieves a sufficiently high improvement in the population, and indeed is a worthwhile investment. It is then used to design a sufficiently powered clinical trial to show this benefit. The developed simulation software is available as open source. We anticipate that this approach can be applied to many other drugs where the benefit or optimal modalities of implementing MIPD is uncertain.

7.3. Methods

7.3.1. Source data

A retrospective study[111] of 315 kidney allograft recipients transplanted between 2004 and 2014 was repurposed for this work. For an in-depth description, we refer to the work by Vanhove et al. The dataset consisted of trough concentrations measured on days 0 to 14 post transplantation under standard of care, i.e. leading to a dose adaptation by experienced transplant physicians targeting trough concentrations between 12 and 15 ng/mL. Extensive data management was required to prepare this dataset for modeling (described in Supplementary Materials). Missing covariates were imputed as the population median.

7.3.2. Model development

Based on literature review, a 1-compartment model with oral absorption was selected as appropriate starting point for this sparse dataset. Following the

[111] Vanhove et al., "Pretransplant 4β-Hydroxycholesterol Does Not Predict Tacrolimus Exposure or Dose Requirements During the First Days After Kidney Transplantation".

approach from other studies,[112] the absorption rate constant (ka) was fixed to $4.5h^{-1}$,[113] as it cannot be reliably estimated from trough data alone. A 2-compartment model and addition of lag time were investigated as possible improvements. The use of a hematocrit-standardized model[114] was investigated including concentration-dependent binding of tacrolimus to erythrocytes. The whole-blood concentration C_{wb} is related to hematocrit-standardized concentration C_{std} through a concentration-dependent proportionality factor R, with C_{stdmax} reflecting the maximum binding capacity, and C_{std50} the concentration associated with half maximum binding.

$$C_{wb} \approx C_b = C_{std} * R * \frac{Hct}{45\%} \qquad (2)$$

$$R = C_{stdmax} * \frac{C_{std}}{C_{std} + C_{std50}} \qquad (3)$$

As precision dosing targeting C_{wb} depends on future hematocrit values, a joint model was used predicting both tacrolimus whole-blood concentration and hematocrit. The time course of hematocrit was modeled using a sigmoid model. IIV was applied to all parameters.

$$Hct = Baseline_{Hct} - EMax_{Hct} * \frac{t}{t + E50_{Hct}} \qquad (4)$$

Random effects were modeled using lognormal inter-individual variability (IIV). Exploratory graphical analysis pointed to a potential increase in clearance over time. This was estimated using inter-occasion variability (IOV) and investigated for correlation with available covariates and time

---

[112] Antignac et al., "Population Pharmacokinetics and Bioavailability of Tacrolimus in Kidney Transplant Patients"; Velickovic-Radovanovic et al., "Population Pharmacokinetics of Tacrolimus in Kidney Transplant Patients"; Han et al., "Prediction of the Tacrolimus Population Pharmacokinetic Parameters According to Cyp3a5 Genotype and Clinical Factors Using NONMEM in Adult Kidney Transplant Recipients".

[113] Jusko et al., "Pharmacokinetics of Tacrolimus in Liver Transplant Patients*".

[114] Størset et al., "Importance of Hematocrit for a Tacrolimus Target Concentration Strategy".

since transplant. The effect of time was modeled as either an exponential (Eq. (5)) or sigmoidal (Eq. (6)) function, with $T50$ the time at which 50% of maximum clearance was reached.

$$CL/F = CL0/F * \left(1 - e^{\frac{ln(0.5)}{T50}t}\right) \qquad (5)$$

$$CL/F = CL0/F * \frac{t}{T50 + t} \qquad (6)$$

Covariates were selected using a stepwise covariate search, including covariates that improved the OFV by 3.84 or more ($p < 0.05$) in the forward step, and eliminating covariates resulting in less than 7.88 increase in OFV ($p > 0.005$). Continuous covariates (age, weight) were included as a power-model (Eq. (7)). Discrete covariates were included as an x-fold change (Eq. (8)). Last observation carry forward (LOCF) was used to interpolate time-varying covariates.

$$X = \theta_X * e^{\eta_X} * (COV/\mu_{COV})^{\beta_{X,COV}} \qquad (7)$$

$$X = \theta_X * e^{\eta_X} * \left(\beta_{X,COV}\right)^{COV} \qquad (8)$$

Only covariates available in routine clinical practice were considered for covariate building. This was therefore limited to age, bodyweight, hematocrit and formulation. Notably, *CYP3A5* genotype was excluded from covariate search, as it is not routinely measured in clinical practice.

7.3.3. Model evaluation

Models were evaluated using the likelihood ratio test, with a ΔOFV of -3.84 or more justifying the addition of a new parameter at $p<0.05$. An evaluation of prediction-corrected visual predictive check (pcVPC), goodness of fit plots, and biological plausibility of parameters was also performed, rejecting models with qualitatively poor results. Prediction-corrected VPCs were generated from 500 simulated subsets of the original data. This was used to identify the appropriate model structure, covariates, inter-individual variability and residual error model.

Models with good population fit were further evaluated on fitness for use in Bayesian feedback for MIPD. To this end, prospective evaluation with Bayesian feedback was used to evaluate predictive performance. Prediction error for observation $n + 1$ was calculated from an individual parameter

estimate on observations $1..n$. Only cases with n and $n+1$ on consecutive days were included.

$$PE\% = \frac{IPRED - CONC}{CONC} \qquad (9)$$

Prediction error $PE\%$ is based on the model prediction $IPRED$ and the actual measured concentration $CONC$. It was used to characterize predictive performance of candidate models as root mean squared error (RMSE), and models were compared using a t-test on $(PE\%)^2$. This is especially relevant to assess fitness for use in MIPD, as prediction error $PE\%$ is directly related to the error in the resulting concentration C_{res} after applying the recommended MIPD dose targeting C_{target}. The intermediate steps are available in Supplementary Material.

$$PE\% = \frac{C_{target} - C_{res}}{C_{res}} \qquad (10)$$

Based on equation (10), an allowed C_{res} between 12 and 15 ng/mL, and C_{target} of 13.5 ng/mL resulted in an allowed $PE\%$ between $[-10\%, 12.5\%]$. This was used to derive a theoretical upper limit for MIPD PTA. We assumed an unlimited number of concentrations, allowing to identify individual parameters but not predict future within-subject variability.

$$\lim_{n->\infty} PTA = p\langle \mathcal{N}(0, \sigma^2) \in [-10\%, 12.5\%] \rangle$$

To allow comparison, standard dosing by physicians was also characterized in the form of $PE\%$, by assuming physicians prescribed a dose that they think will hit the target. In other words, the *in cerebro modeling* of the physician predicts a trough concentration of 13.5ng/mL at the dose they prescribed. The $PE\%$ for physicians was therefore calculated as

$$PE\% = \frac{13.5 - CONC}{CONC} \qquad (11)$$

7.3.4. Bayesian feedback and model-predictive control

Tacrolimus exhibits high PK variability, not only between patients but also within a single subject. This variability can be described by the residual error model or through inter-occasion variability, depending on whether rich individual data is available. To the best of our knowledge, all models previously used to describe tacrolimus pharmacokinetics assume a random variability. This variability is not entirely random however, and previous

studies have reported a correlation of the error between subsequent occasions[115]. This seems intuitive: an increased clearance on day N should indeed carry over to the following day N+1. This was quantified in the dataset by calculating the autocorrelation of the residual error from a standard bayesian fit. For a detailed discussion, please see Supplementary Material.

To integrate this correlation between subsequent occasions, we opted to use a pragmatic approach, inspired by similar closed loop control systems in anesthesia.[116] On day 1, regular empirical bayesian estimation (EBE) with a priori estimates Θ and inter-individual variability Ω is used to estimate the most likely individual parameters η. On day i, the individual parameter estimates from day $i-1$ are used as *a priori* estimates ($\Theta' = \eta$), while the same inter-individual variability Ω is retained. This approach is reminiscent of model-predictive bioreactor control in chemical engineering and was therefore dubbed model-predictive control MIPD (MPC/MIPD). The predictive performance of this method was compared to classical EBE using the prediction error $PE\%$ described previously.

115 Størset et al., "Improved Prediction of Tacrolimus Concentrations Early After Kidney Transplantation Using Theory-Based Pharmacokinetic Modelling".

116 Krieger and Pistikopoulos, "Model Predictive Control of Anesthesia Under Uncertainty".

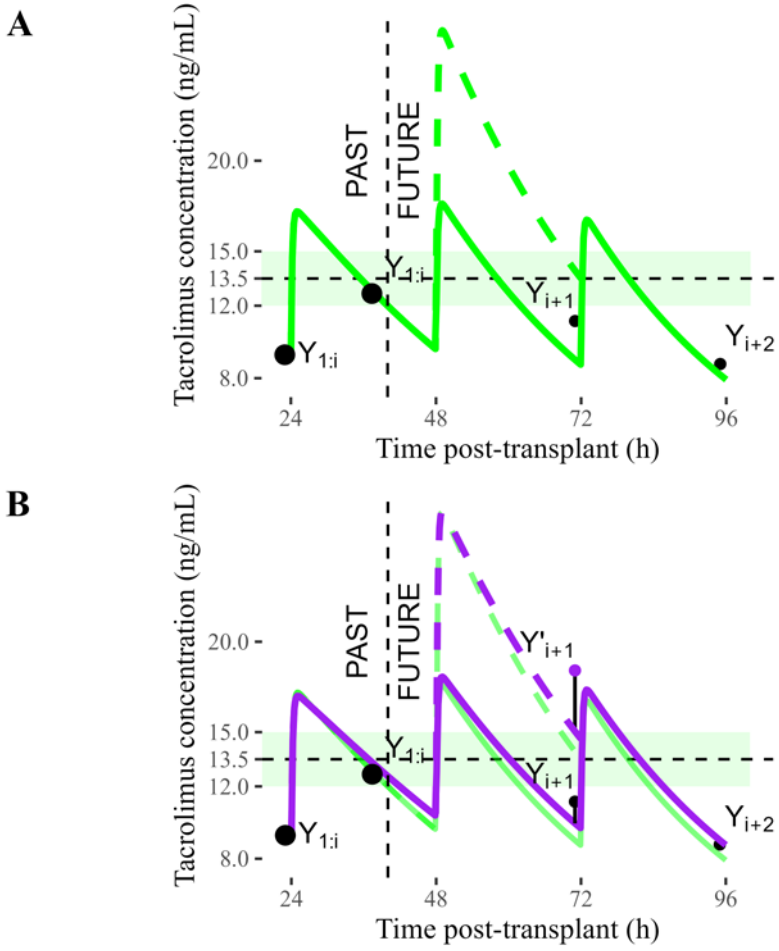


***Figure 25**: (A) Observations 1 to i were used to calculate fit $f(\eta)$ (solid green line). This was then used to find dose D_rec that makes $f(D_{rec}, t_{j+1}) = C_{target}$ (dotted green line). (B) All observations 1:n were used to calculate fit $f(\eta_{orig})$ (solid purple line). This was then used to find $f(\eta_{orig}, D_{rec}, t_{i+1})$ (dotted purple line). The final result Y'_{i+1} is calculated by adding the original residual error ϵ on top.*

7.3.5. Dosing algorithm and simulation strategy

Once a method for predicting future concentrations was established, we could search the dose required to optimize resulting concentrations. The dosing algorithm is described by the following pseudo-code, with Y_i the

measured concentration, t_i the associated sample time, η the individual parameter estimates, D_j the dose at administration j, t_j the associated dosing time, t_{j+1} the time of the subsequent dose, $f(\eta, t_{j+1}, D_j)$ the function to predict a concentration at time t_{j+1} as a result of all doses up to and including dose j, and C_{target} the target concentration:

```
INITIALIZATION:
  observed = []
  regimen = [ Loading Dose, Future Planned Doses]
WHEN A NEW CONCENTRATION (t_i, Y_i) BECOMES AVAILABLE:
  add (t_i, Y_i) to observed
  update regimen from the patient EHR system
  fit η using MPC/MIPD
  FOR all future doses at time t_j:
    use a root finding algorithm to find dose D_rec,j
    such that:
        f(η, D_rec,j, t_j+1) = C_target
    and adapt the regimen: D_j = D_rec,j
  if simulating:
  determine the next observed concentration
    Y_i+1' := f(η_orig, t_i+1) + ϵ
```

The dosing algorithm was designed to execute as new concentration samples become available, regardless of whether that sample is within the acceptable range. The full dosing history and concentrations at each iteration were used

to fit individual parameters η, after which the PK model $f(\eta, D_{rec,j}, t_{j+1})$ was used to predict future trough concentrations. For each future administration, the dose was adapted such that the trough concentration was equal to the target concentration C_{target} of 13.5ng/mL. This is visualized in Figure **25**a.

We made sure computer dosing would require minimal changes to the current clinical workflow at UZ Leuven. Therefore, the loading dose was not adapted. Only dosing amounts were adapted. Planned dosing times and formulations (Advagraf® or Prograft®) were retained from the source dataset. Doses of 0mg were not recorded in the source dataset, therefore a dummy dose of 0mg was added at 08:00 if Advagraf® was previously administered, or 08:00 and 20:00 if Prograft® was previously administered. To ensure fair comparison with physician-based dosing, only doses 6 hours after a concentration sample were considered for adaptation, as concentrations were generally only available in practice 3 hours post-sampling and adapting a dose close to administration time was not deemed practical. Doses were rounded to 0.5mg.

To simulate the resulting concentrations Y' after applying MIPD, the following procedure was used (visualized in Figure 1b). As we cannot go back in time and administer the recommended dose D_{rec} to the actual patient, we used the best prediction available: we simulated using a fitted η_{orig} on all observed concentrations for this patient in the historic dataset $f(\eta_{orig}, t_{i+1})$ and re-applied the original residual error ϵ.

$$Y'_{i+1} = f(\eta_{orig}, t_{i+1}) + \epsilon_{i+1} \qquad (12)$$

This exercise was performed on all 315 patients for all available trough samples. Missing concentration samples were reused as missing data. This resulted in two parallel datasets: a first dataset of dose and resulting concentration per patient per day as originally performed in reality by physicians in standard of care, and a second hypothetical dataset where the dose was calculated through MIPD. Both arms could then be compared graphically and using statistical methods.

In collaboration with physicians, an *improvement* by MIPD was qualitatively defined as: more patients with trough concentrations in the target window of 12 to 15 ng/mL, faster target attainment, and smaller deviations from target. This was quantitatively defined as (i) higher individual probability of target attainment, (ii) faster attainment of 1, 2, 3, … cumulative days in target, and (iii) smaller overall distance to target on each day, defined in equation (13)

$$C > 15 \Rightarrow D^2 = (\log C - \log 15)^2 \qquad (13)$$

$$C < 12 \Rightarrow D^2 = (\log C - \log 12)^2$$

$$C \in [12,15] \Rightarrow D^2 = 0$$

7.3.6. Statistical methods and power calculation

Based on the above population simulations, a good description of PK outcome for N=315 individuals was available. This allowed defining statistical tests to quantify the effects in the population. Furthermore, a power calculation was performed to consider what effect size could be significantly proven in a trial.

Dose adaptation performance was expected to be time dependent. Any closed loop system requires some samples to reach the target, and overshoot, undershoot or 'lucky hits' are to be expected. The proposed statistical analysis accounted for these effects.

1. Improvement on individual PTA was evaluated as a Welch's t-test. A relative improvement of +33% was deemed clinically relevant.
2. Speed of target attainment was evaluated as a time-to-event process with non-proportional hazards. A one-sided Mantel-Haenszel log-rank test on TTE >3 concentrations in target was used. A minimum relative improvement of +33% fraction of patients reaching target on day 7 was deemed clinically relevant. Power for this test was calculated based on required difference in relative hazard ratio and the expected events over the accrual period of 14 days. (see Supplementary Materials for more details)
3. We expected squared log-distance to target to decrease over time. Ideally, a mixed model repeated measurements (MMRM) model detects a significant reduction of squared log-distance due to MIPD. Power for this test was not calculated, as no established method for power analysis of MMRM models with non-normal outcomes is available as of yet.

Based on the simulation, accurate estimates of the distribution of these statistics were available. These were subsequently used to determine required sampled size to detect clinically relevant effect.

7.3.7. Clinical trial simulation

Finally, the candidate trial with N=200 patients at 2:1 allocation was evaluated as a clinical trial simulation. Standard of care was not simulated,

but rather sampled without replacement from the available N=315 profiles in the retrospective dataset. The MIPD arm was similarly sampled from the profiles in the simulation previously performed. Dropout and missing data were considered as represented realistically in the retrospective dataset. This was repeated 1000 times to characterize the distribution of possible clinical trial outcomes and evaluate Probability of Study Success (PoSS).

7.3.8. Software

Monolix 2019[117] was used to perform modeling, using the SAEM algorithm complemented with importance resampling to determine -2 log-likelihood. R version 3.5.2 was used for all data management and simulation tasks, using tdmore version 1.1.[118] Tdmore is under active development and can be freely downloaded.

[117] "Monolix Version 2019r1".

[118] Faelens, Luyckx, and Quentin Leirens, "Tdmore".

***Table 3**: Parameter estimates for hematocrit-standardized, base and full model.*

Parameter	Joint model (OFV=1594.6*)	RSE	Base model (OFV=19669.68)	RSE	Full model (OFV=19560.72)	RSE
Typical values						
Ka [/h]	4.5	fix	4.5	fix	4.5	fix
V [L]	562	2.9%	767	3.2%	760	3.1%
CL0 [L/h]	17.8	2.5%	27.6	2.6%	27.2	2.5%
T50 [h]	19.5	6.7%	26.4	5.6%	25.7	6.3%
Hct baseline [%]	0.467	0.49%				
Hct Emax [%]	0.188	0.71%				
Hct T50 [h]	1.23	20%				
Covariate effects						
Haematocrit on CL					-0.461	0.33%
Weight on CL					0.571	NaN%
Weight on V					0.536	0.21%
Inter-individual variability						
V	57.4%	4.8%	62.8%	4.9%	60.6%	5%
CL0	53.8%	4.2%	55.7%	4.2%	52.9%	4.3%
T50	136%	6.9%	117%	6.2%	123%	6.8%
Hct baseline	7.56%	4.4%				
Hct Emax	5.7%	11%				
Hct T50	710%	7.8%				
corr V,CL0	0.715	4.6%	0.671	5.5%	0.68	5.5%
corr HctT50,HctBaseline	-0.69	8.2%				
corr T50,HctBaseline	-0.375	17%				
corr T50,HctT50	0.5	15%				
Residual error						
Proportional	0.183	1.3%	0.187	1.3%	0.185	1.3%

Inter-individual CV% was calculated as exp(omega)-1. Relative standard error (RSE) was determined through importance resampling. RSE for the effect of bodyweight on CL could not be determined numerically. *: includes hematocrit observations

7.4. Results

7.4.1. Model building

Base model Key model parameter estimates are available in Table **3**. Relevant diagnostic plots are available in Supplementary Material. A 1-compartment model with oral absorption showed considerable time-dependent bias on individual weighted residual (IWRES) vs time plots, overpredicting early (before day 4) concentrations and underpredicting late concentrations (days 7 and later). Inclusion of hematocrit-normalized concentration improved the fit ($\Delta OFV = -356.83$), but did not reduce time-dependent bias. Concentration-dependent binding was removed without any notable impact ($\Delta OFV = 0.14$). Estimation of IOV on clearance further showed a time-dependent trend, which was most appropriately modeled through Eq. (5), yielding a T50 of 38.7h and $\Delta OFV = -1875$. The pcVPC showed acceptable fit on median; the outer prediction interval improved by adding IIV on T50 ($\omega_{T50} = 125CV\%, \Delta OFV = -522$) and correlation between η_V and η_{CL} ($\rho = 0.681, \Delta OFV = -145.57$). Additive residual error was subsequently removed without impact to OFV or fit. The model did not further improve through absorption lag time or a 2-compartment disposition.

Predictive performance Prospective evaluation showed high prediction error (RMSE of 0.361) due to bias introduced by using LOCF for hematocrit. Joint modeling of tacrolimus and hematocrit through Eq. (4) retained good population fit and greatly improved predictive performance (RMSE of 0.307, p-value 0.004). By moving from 3 to 6 estimated individual parameters, simulation time increased 40-fold. For further simulations, the hematocrit-standardized model was removed. This increased OFV by 213 points but did not significantly decrease predictive performance (RMSE of 0.325, p=0.172).

Covariate search Stepwise covariate modeling (SCM) is described in Table S1. As significant covariates, we identified hematocrit on clearance ($\Delta OFV = -79.65$), bodyweight on clearance ($\Delta OFV = -13.48$) and age on clearance ($\Delta OFV = -10.04$).

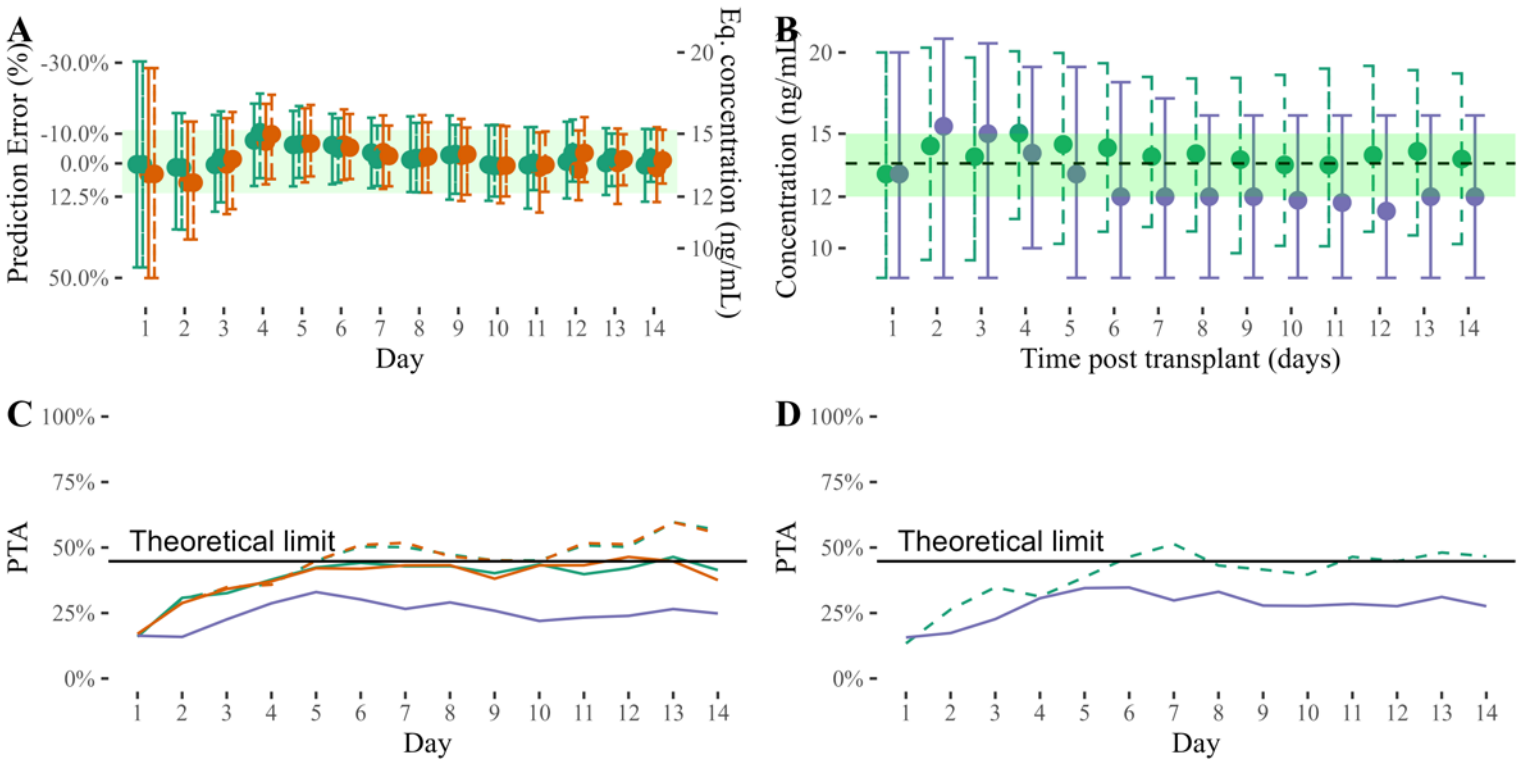


***Figure 26**: Relative prediction error (A) and predicted concentrations (B) median and 50% prediction interval for physician (blue), base (green), and full (orange) model, using EBE (solid line) and MPC (dotted line). The target window is represented as green area. Probability of target attainment is shown for the relative prediction error (C) and predicted concentration (D). The theoretical limit is derived from the model residual error.*

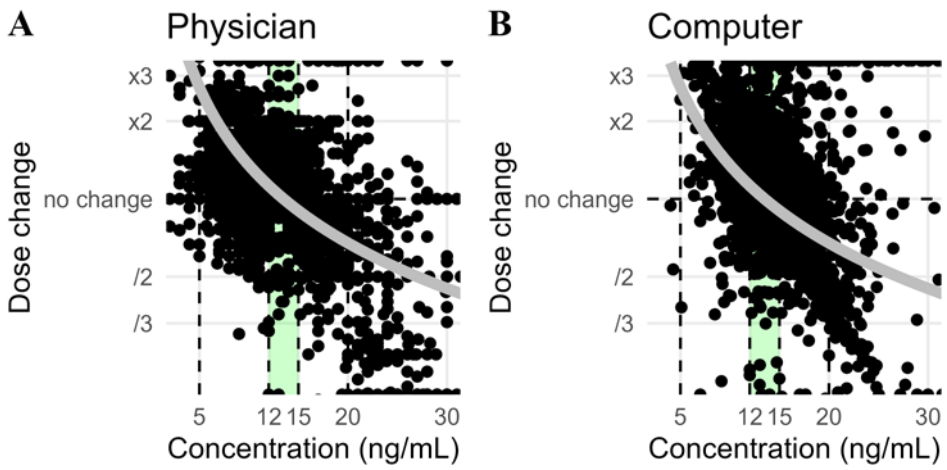


***Figure 27**: Physician dose adaptation (left) versus computer dose adaptation (right). The observed concentration (x-axis) results in a dose change (y-axis). The grey line shows the theoretical dose adaptation when following the rule of three in steady-state.*

7.4.2. Model-predictive control

Autocorrelation of the base model residual error was $E[\rho_k] = 0.52$, 0.37, 0.24, 0.16, 0.13, 0.12, 0.10, and 0.12 for k-values of 1 (autocorrelation between subsequent observations) to 8 (autocorrelation between observations 8 days apart). This points to correlated consecutive residual errors. Predictive performance of MPC/MIPD is shown in Figure **26**A. Compared to EBE, MPC/MIPD shows a significant improvement in predictive performance. This applies to both base and full model. Using these results, the base model with MPC/MIPD estimation was selected as the optimal approach, at RMSE of 0.304 (p=0.432 versus hematocrit-standardized model, p=0.148 versus base model with EBE estimation).

Based on the identified residual error, the theoretical upper limit for target attainment is 45.2%. Physician performance averaged 25% PTA, with clear underdosing visible in Figure **26**B. For bayesian estimation, PTA (Figure **26**C) approached the theoretical upper limit on six out of fourteen days, while MPC/MIPD exceeded this limit. There is no apparent bias visible in Figure **26**A, as mean prediction error is close to 0. The full model did not outperform the base model. We opted to use the base model in further simulations, as collecting covariates was not worth the increased clinical workload.

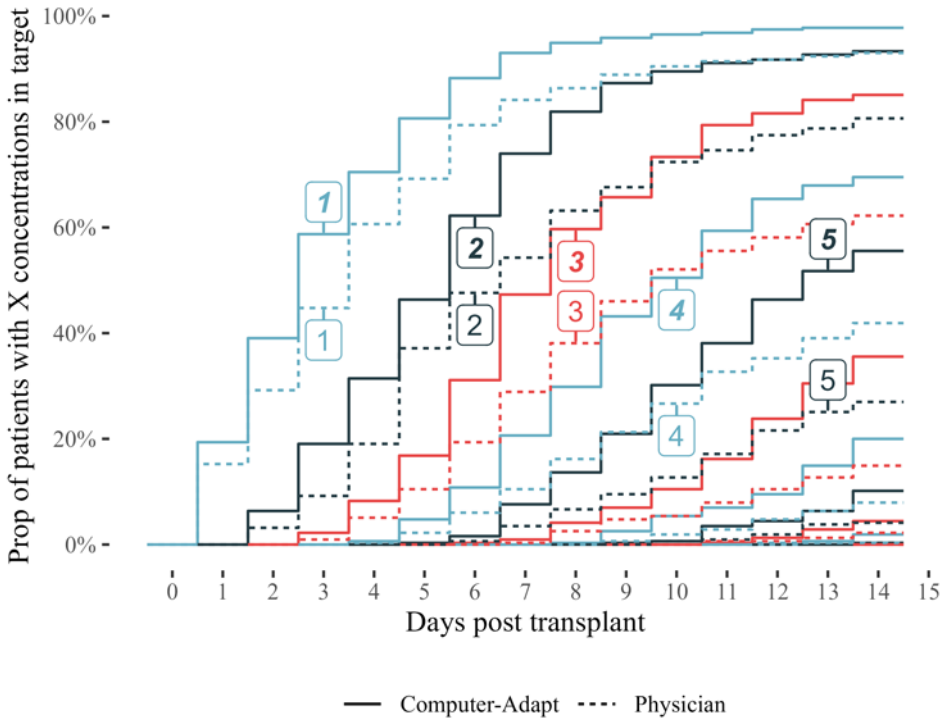


***Figure 28**: Proportion of patients with at least X concentrations in target, per day. Computer (solid line, bold font) vs physician (dotted line, normal font).*

7.4.3. Simulation of model-informed precision dosing

The base model was used with MPC/MIPD to simulate dose adaptation and resulting concentrations. The dose adaptations performed by physicians and MIPD are compared in Figure **27**. While physicians adapted conservatively, MIPD applied a temporary overcorrection of the dose in order to reach target concentration as fast as possible. Figure **26** shows a summary of concentration per day. MIPD resulted in a large PTA, as well as overall concentrations closer to the target window. Per-patient PTA was at 39% ± 15.8% for MIPD (mean ± standard deviation), while physician PTA was at 28% ± 16.1%. Time-to-event curves for 'X observations in the target window' are shown in Figure **28**. The difference for reaching ">1 day in target" is quite small, with only a one-day delay between arms on average. This delay grows larger, with ">3 days in target" being reached for 50% of the population on day 8 for the intervention arm, while only at day 10 for the

control arm. On day 7, 48.2% ± 5.5% reached >3 days in target for MIPD, while only 27.5% ± 3.51% reached this for the physician arm. Finally, the KS-test identified a significant reduction in squared log-distance to target window for every day. Log-squared distance to target was normally distributed after Box-cox transformation, allowing the application of an MMRM analysis. This identified a significant treatment effect, yet only at a relative improvement of -13%. Squared log-distance to target window decreased from 0.080 ± 0.202 to 0.055 ± 0.191 (mean ± standard deviation).

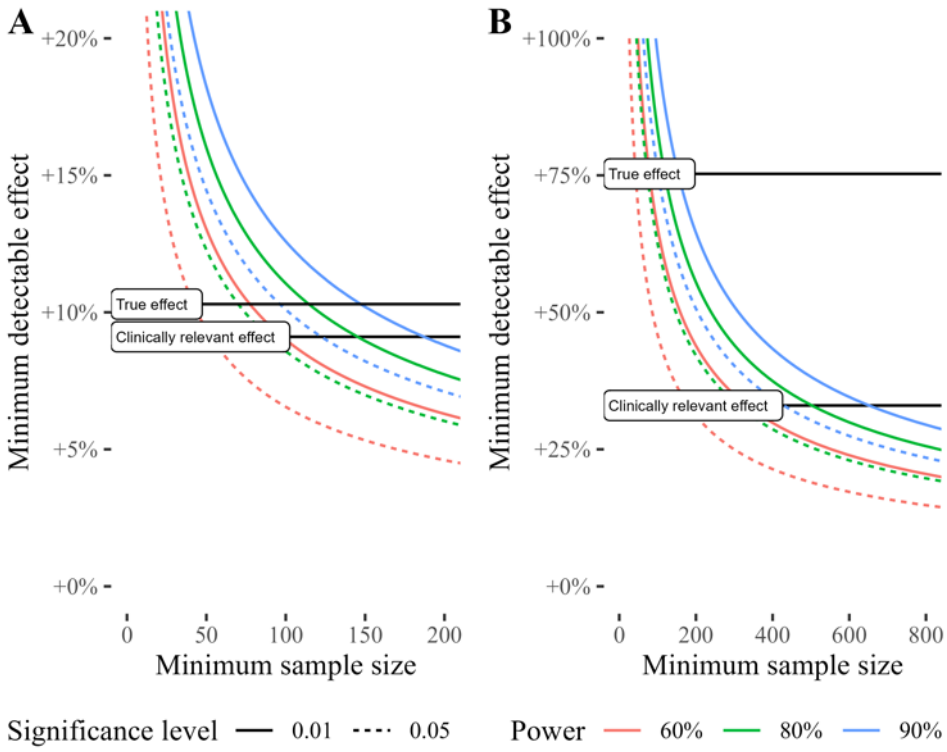


***Figure 29**: Minimum detectable effect size for difference in PTA (A) and difference in speed of reaching 3 concentrations in target (B). Horizontal lines show the true effect and minimum clinically relevant effect*

7.4.4. Power calculation

Based on the above estimates, study power and minimum detectable effect sizes are presented in Figure **29**. The candidate trial of N=200 will reliably detect a PTA improvement at p<0.01. The clinically relevant PTA can be

detected at p<0.01 with N=145 patients. An improvement on TTE less than +50% may not be reliably detected by a trial with N=200 patients, yet the true effect will be detected even at p<0.01. On log-squared distance to target, the population simulation showed a true effect size below the clinically relevant limit.

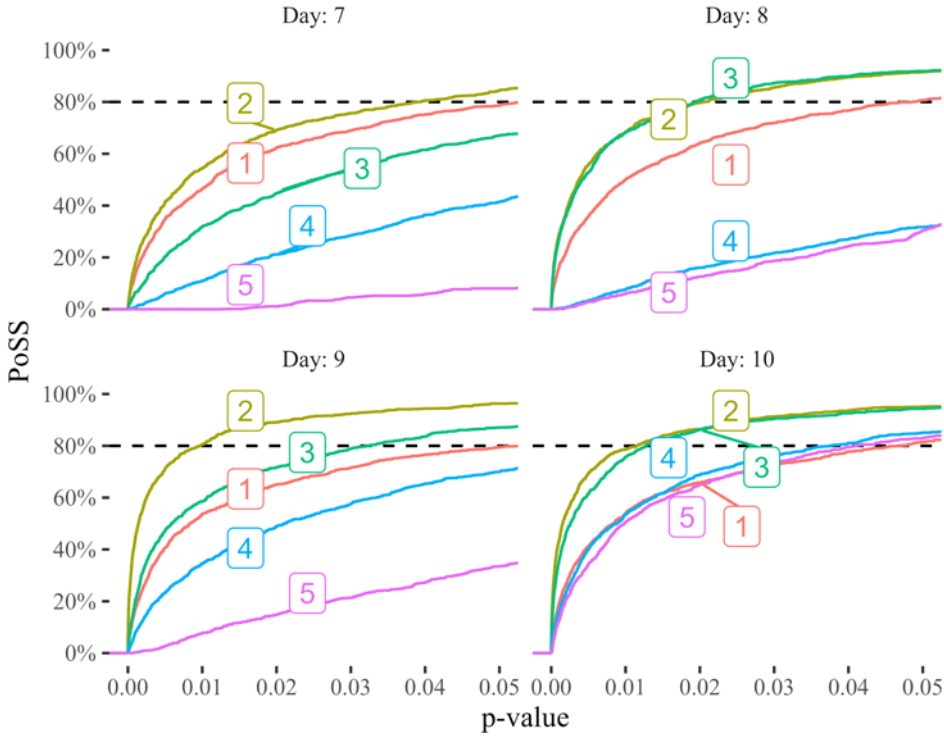


***Figure 30**: Power to detect an improvement in "time to reach >X concentrations in target" (X=1 to 5) after 7, 8, 9 or 10 days, using Mantel-Haenszel test. PoSS: probability of study success.*

7.4.5. Clinical trial simulation

Study power to detect these three aforementioned endpoints was simulated using a bootstrap of 1000 random trials of 200 patients each. It was trivial to show an improvement in average PTA per patient, with a 100% probability of study success (PoSS) at $p < 0.01$. The average expected effect size was 11.3% [8% - 14%], with per-trial 95% lower confidence limit of 7.5% [4% - 10.5%]. Study power for the TTE test is shown in **30**. PoSS depended on the

day and the endpoint both arms were compared at. It was decided to evaluate "3-days-in-target" at day 8 post-transplant, which yielded a PoSS of 90% at $p < 0.01$. Looking at squared-log-distance-to-target, even the sensitive KS-test could only identify an improvement on days 3, 7, 8 and 10 post-transplant. Only at these days was PoSS > 80% for $p < 0.01$. Using MMRM analysis, an overall improvement could be reliably shown, although per-day effects could only be reliably shown on days 6, 7, 8 and 11 at $p < 0.05$.

7.5. Discussion

To the best of our knowledge, this is the first example of predicting a clinical trial outcome comparing model-informed precision dosing to standard of care, and the use of this prediction to optimize a future planned prospective clinical trial. To achieve this, a population PK model was first built and evaluated for goodness of fit on the target population. A pragmatic approach was presented to incorporate unexplained variability in individual parameters. The full simulation code was implemented in a reusable R package. Finally, the predicted results were analyzed to describe the statistical power of a candidate clinical trial design, allowing optimization of said trial design.

Overall, the population PK model is in reasonable agreement with literature. Describing tacrolimus PK by a 1-compartment model with oral absorption is common in the absence of rich concentration-time profiles.[119] Independent groups identified similar time-dependent clearance early post-transplant.[120]. Others identified a time-dependent increase over several weeks post-transplant[121], which is likely a different effect altogether. Parameter

[119] Campagne, Mager, and Tornatore, "Population Pharmacokinetics of Tacrolimus in Transplant Recipients".

[120] Antignac et al., "Population Pharmacokinetics and Bioavailability of Tacrolimus in Kidney Transplant Patients"; Han et al., "Population Pharmacokinetic-Pharmacogenetic Model of Tacrolimus in the Early Period After Kidney Transplantation"; Zuo et al., "Effects of Cyp3a4 and Cyp3a5 Polymorphisms on Tacrolimus Pharmacokinetics in Chinese Adult Renal Transplant Recipients".

[121] Han et al., "Prediction of the Tacrolimus Population Pharmacokinetic Parameters According to Cyp3a5 Genotype and Clinical Factors Using

estimates are broadly in agreement with results from similar studies focusing on the first 14 days post transplant[122]. Identified covariates are also in agreement with previous studies[123], although some identified a large difference in bioavailability between Prograft® and Advagraf® formulations. The identified power factor $\beta_{CL,WT}$ of 0.348 results in a 77% and 122% adjustment of clearance for the lightest and heaviest patients (33.5kg and 125kg respectively) in the study, which in light of inter-individual variability of +- 55.7% explains the minimal difference between base and full model predictive performance. The identified IIV is high, which we attribute to the poor PK stability of patients early post-transplant. This agrees with other studies focusing on the same study period.[124]

Notably, the inclusion of hematocrit at first resulted in poor predictive performance. When the full profile of a time-varying covariate is not available during prospective evaluation, significant bias may be introduced. Joint modeling of both drug concentration and covariate is required to overcome this limitation. This markedly increased computation times. While model simplification resulted in a penalty to OFV, predictive performance was not significantly impacted. Notably, applying MPC/MIPD again resulted in low

---

NONMEM in Adult Kidney Transplant Recipients"; Størset et al., "Improved Prediction of Tacrolimus Concentrations Early After Kidney Transplantation Using Theory-Based Pharmacokinetic Modelling"; Bergmann et al., "Population Pharmacokinetics of Tacrolimus in Adult Kidney Transplant Patients"; Golubović et al., "Total Plasma Protein Effect on Tacrolimus Elimination in Kidney Transplant Patients – Population Pharmacokinetic Approach".

122 Antignac et al., "Population Pharmacokinetics and Bioavailability of Tacrolimus in Kidney Transplant Patients".

123 Campagne, Mager, and Tornatore, "Population Pharmacokinetics of Tacrolimus in Transplant Recipients".

124 Antignac et al., "Population Pharmacokinetics and Bioavailability of Tacrolimus in Kidney Transplant Patients"; Staatz and Tett, "Clinical Pharmacokinetics and Pharmacodynamics of Tacrolimus in Solid Organ Transplantation"; Musuamba et al., "Statistical Tools for Dose Individualization of Mycophenolic Acid and Tacrolimus Co-Administered During the First Month After Renal Transplantation".

RMSE, rivaling the more complex model with more covariates, at feasible calculation times.

McDougall et al[125] extensively explored the impact of model misspecification on precision dosing. They demonstrated that only severe model misspecification significantly impacts model-based precision dosing performance. This reasoning also applies to covariate models; covariates difficult to collect can be omitted without impact to model predictive performance. This further exemplifies the necessity to include prospective evaluation in the diagnostic toolset when developing models for precision dosing.

Tacrolimus PK has typically been described by a 2-compartment model when rich data are used[126]. In this case however, the use of 2-compartment kinetics would not result in different results. The typical distribution phase is $< 12h$, and therefore no information on the distribution phase is present in daily trough concentrations, even with multiple dosing. There is an ongoing debate on appropriate PK targets for tacrolimus, with some evidence pointing to AUC as a superior metric[127]. Our trough dataset cannot be used to accurately predict AUC[128]. With rich data and an appropriate model, the presented approach may be applied to evaluate the accuracy (and improvement over standard of care) when using 1, 2, or more blood samples per day.

CYP3A5 genotype is missing from our current model, as it was not generally measured in transplant patients at UZ Leuven hospital. It is worthwhile to evaluate the inclusion of this covariate, and to quantify the potential improvement in MIPD dosing accuracy. If this covariate is not available, a mixture model could be used to estimate individual CYP3A5 expression

[125] McDougall et al., "The Impact of Model-Misspecification on Model Based Personalised Dosing".

[126] Brooks et al., "Population Pharmacokinetic Modelling and Bayesian Estimation of Tacrolimus Exposure".

[127] Brunet et al., "Therapeutic Drug Monitoring of Tacrolimus-Personalized Therapy".

[128] Op den Buijsch et al., "Evaluation of Limited Sampling Strategies for Tacrolimus".

probability. However, due to the low number of CYP3A5 expressors in the target population, we expect the impact to be low and transient.

In this work, we argue that classical goodness of fit evaluation is not appropriate when building a model for MIPD. Even though the full model is a significantly better description of the data as compared to the base model, this did not result in a significant improvement to predictive performance or PTA. Clinically, it is preferable to omit covariates that are cumbersome to measure, if they do not improve predictive performance significantly. In general, we argue that predictive performance assessment is a key step when building MIPD models.

However, contrary to the well-studied classical goodness of fit evaluation, it is unclear how a model with poor predictive performance can be improved. As a first step, we suggest to include predictive performance evaluation in model building software. The Perl-Speaks-Nonmem suite recently added the `proseval` tool, but lacks clear standard graphs to represent this data.

The MPC/MIPD method merits further discussion. Correlated residual errors in tacrolimus models were previously identified by Størset et al,[129] who reported that bioavailability varied less between subsequent occasions. Correlated inter-occasion variability, which could otherwise be classified as 'parameter drift', has not been studied in detail. Pragmatic solutions include down-weighing older concentration samples or arbitrarily increasing ω during estimation. The novel idea of adapting the estimation method rather than the model resulted in a significant improvement to predictive performance in this dataset.

The predicted outcome, in the form of predicted tacrolimus trough concentrations for the MIPD arm, differs from the estimated model predictive performance for three reasons. Firstly, an accurate prediction does not necessarily imply a future concentration in target. We can use the first dose recommendation as an example: the loading dose is too high in 170 out of 315 patients, and using the first trough concentration at 08:00 on day 1, the computer recommends a 0mg evening dose. The computer predicts this dose will still result in too high concentrations the following morning. Secondly, the opportunity to adapt the dose may be far into the future, challenging the

[129] Størset et al., "Improved Prediction of Tacrolimus Concentrations Early After Kidney Transplantation Using Theory-Based Pharmacokinetic Modelling".

predictive performance of the model under high within-patient variability. When the dose adaptation is performed at 12:00, the Prograft® administrations of 20:00 evening and 08:00 the following morning can be adapted, with the latest trough therefore at 20:00 the next day. For Advagraf however, only the dose at 08:00 the following morning can be adapted, with trough at 08:00 two days into the future. This increases prediction error and therefore reduced probability of target attainment. Finally, the prediction error does not directly translate to an error in resulting concentration after MIPD for non-steady state. This highlights the importance of using PK models for dose adaptation: dose adaptation tables using dose-normalized concentration fail to capture the highly variable and non-steady-state nature of the first 2 weeks post transplant.

Figure **27** shows a computer algorithm performs aggressive dose adaptation, in stark contrast to conservative dose adaptation by physicians, who seem to be more cautious in this respect. We offer three explanations for this behavior. Firstly, *in cerebro* modeling assumes steady state, and therefore cannot correctly relate a wildly varying dosing history and concentrations to the required dose adaptation. Secondly, it is difficult for humans to capture pharmacokinetic dose-linearity. If the concentration is 50% below target, the dose should be doubled. Instead, we see slow uptitration by absolute steps, rather than e.g. doubling or halving the dose, contrary to current research advising against tacrolimus underexposure.[130] Finally, we identified time-dependent clearance during the first week post transplant. Even when gradually increasing the dose, doctors are chasing a moving goalpost. MIPD does not suffer from any of these shortcomings.

In contrast, MIPD even employs a corrective dose to ensure the target trough concentration is reached as soon as possible. It remains unclear whether this practice is beneficial in real life. Firstly, there is an ongoing discussion on the validity of trough concentration as a PK target.[131] Targeting AUC may be

[130] Brunet et al., "Therapeutic Drug Monitoring of Tacrolimus-Personalized Therapy".

[131] Bouamar et al., "Tacrolimus Predose Concentrations Do Not Predict the Risk of Acute Rejection After Renal Transplantation".

more appropriate.[132] A recent study by Miano et al[133] identified a PK/PD association for safety. A 54% increase in acute kidney injury was identified per 5ng/mL increase in average tacrolimus trough concentrations over the previous 3 days. A similar association for efficacy could not be identified. Adding clinical utility (CU), a model integrating PK/PD/CU could focus on true benefit for patients, rather than improvement on surrogate endpoints with only weak association to clinical benefit.

In this work, we demonstrated clearly that *simulated MIPD concentrations* can serve to refine the definition of trial endpoints. This allowed us to explore beyond mere "improvement in average PTA" and evaluate endpoints such as "speed of target attainment" and "distance to target window". It was not possible to detect an improvement consistently on each separate day, even using advanced statistical techniques such as Mixed Model Repeated Measurements (MMRM). Only a consistent effect across all days could be shown reliably. All things considered, the proposed techniques dig deeper into MIPD performance than evaluating odds-ratios of PTA. As of yet, the relevance of the presented surrogate endpoints and their relevance to clinical outcome is based on empiric evidence only. It is unfortunate that no PK/PD model predicting clinical outcomes has been identified. Such a model could serve to replace a naïve therapeutic drug monitoring approach targeting a therapeutic window, and instead directly find the appropriate dose to target a desired PD effect reaching optimum efficacy and toxicity. This model may also serve to design a concentration-controlled trial quantifying the clinical impact of MIPD.

A randomized controlled trial is always comparative in nature. Therefore, the presented results cannot easily be translated to other hospitals, as the standard of care differs widely between hospitals. As an example, steroid concomitant therapy was identified to influence tacrolimus PK,[134] but

---

[132] Kuypers et al., "Clinical Efficacy and Toxicity Profile of Tacrolimus and Mycophenolic Acid in Relation to Combined Long-Term Pharmacokinetics in de Novo Renal Allograft Recipients".

[133] Miano et al., "Early Tacrolimus Concentrations After Lung Transplant Are Predicted by Combined Clinical and Genetic Factors and Associated With Acute Kidney Injury".

[134] Press et al., "Explaining Variability in Tacrolimus Pharmacokinetics to Optimize Early Exposure in Adult Kidney Transplant Recipients"; Antignac et

different treatments of high-dose steroids are in use at different hospitals. Furthermore, there is no clear evidence that retrospective data will be similar to standard of care performance in a comparative trial. Standards may have improved with increased experience, and a clinical trial setting may invite physicians to more carefully perform dose adaptation to achieve the target window.

In conclusion, this work offers new insights into the use of simulation to predict and optimize MIPD for tacrolimus dose adaptation. We have shown the validity of predictive performance as a tool for model selection. MPC/MIPD was proposed as a method to incorporate unexplained but autocorrelated inter-occasion variability. Retrospective data was used to fully simulate a hypothetical MIPD arm, which was then used to quantitatively analyze the improvement the technique offers. Finally, the simulated data was used to calculate trial power and optimize said trial. The simulation software was implemented as an open-source R package, allowing to repeat this exercise with any model. By making this software available, we hope quantitative predictions on MIPD become within reach, allowing to identify where this technology can benefit clinical care the most.

al., "Population Pharmacokinetics and Bioavailability of Tacrolimus in Kidney Transplant Patients"; Velickovic-Radovanovic et al., "Population Pharmacokinetics of Tacrolimus in Kidney Transplant Patients".

Chapter 8 Tacrolimus MIPD trial simulation

This chapter explores the clinical trial design of a future tacrolimus MIPD trial in more detail. It was a key report in the design of the prospective trial entitled *"Integrated Model-based Medication Dosing Assist App(lication) in Klinisch Werk Station (KWS). Proof-of-Concept Validation Study: Tacrolimus in Kidney Transplantation"* and its accompanying statistical analysis plan.

8.1. Abstract

Introduction The aim of this in silico study was to predict how model-informed precision dosing (MIPD) impacts target attainment of tacrolimus in de novo kidney transplant patients in the first 14 days post-transplant, as compared to standard of care (SoC) where physicians perform dose adaptation without the use of a computer or statistical model. This difference was evaluated in the population, while the statistical power to detect such a difference was also evaluated for the planned study of N=200 patients.

Methods To represent SoC, a retrospective dataset of 315 patients was evaluated. To represent MIPD, these 315 patients were dose-adapted in silico using a computer algorithm. We evaluated the difference between MIPD and Standard of Care on four aspects: time until a patient has a first trough sample in the target window, per-patient target attainment, proportion of concentrations in target per day, and squared log-distance between trough concentration and target window. Overall effect size was predicted through population simulation, and probability of study success (PoSS) was estimated through bootstrap.

Results

- Difference in time to first-sample-in-target was a feasible endpoint. Population simulations show a +8% improvement after 1 day, and +15% after 2 days. A Mantel-Haenszel test showed MIPD superiority in a simulated clinical trial, with PoSS at 87.2% for $p<0.05$, and 62.8% for $p<0.01$. For time to multiple concentrations in target, PoSS was even higher.

- Per-patient target attainment was a feasible endpoint. Average per-patient target attainment increased from 28% to 39%. A t-test was shown to detect MIPD superiority in the clinical trial, with PoSS at 100% for $p<0.01$.
- A difference in proportion of samples in-target per day was not a feasible endpoint. Although population simulation predicted +11.8% more samples in target for MIPD -ranging per day from +5% (day 4) to +20% (day 7)- a proportions test could only reliably (+80% PoSS) detect this on day 7 and 11.
- A difference in overall squared distance from target window was a feasible endpoint. A generalized least squares model with fixed effects per day, per arm, and interaction day/arm described the data well. A significant improvement for MIPD in all days except days 4 and 5 was observed. An overall difference could be reliably detected at $p<0.01$. Per-day, this difference could only be reliably detected (+80% PoSS) on day 6 to 8, and day 11 at $p<0.05$.
- Squared distance from target window per day was not a feasible endpoint. Population simulation showed a significant decrease on all days for MIPD, although the difference is small on days 4 and 5. In the clinical trial, a Kolmogorov-Smirnov test at $p<0.05$ could only reliably (+80% PoSS) detect this difference on day 3 and days 7 through 11.

Conclusion Faster time-to-target, higher per-patient target attainment and lower overall squared distance from target window were all predicted to be detectable at a high probability, indicating they may be used as a primary endpoint. Unfortunately, none of the analysis methods studied were predicted to detect a higher proportion in target per day or lower distance from target window per day on the first days post-transplant.

8.2. Introduction

This chapter details the simulations performed in preparation of the clinical trial *Dose adaptation of tacrolimus for de novo kidney transplant recipients in the first 14 days post-transplant*, scheduled between end of 2019 and end of 2022.

Tacrolimus is an immuno-suppressor that is used to reduce the risk of graft rejection after solid organ transplant. As with many drugs, it has reduced efficacy at low concentrations, but is potentially toxic at high doses. For

tacrolimus, a trough concentration between 10-15 ng/mL is targeted during the first 3 months post transplant.[135]

Unfortunately, tacrolimus pharmacokinetics are widely variable in the population. A standard dose of 5 mg/day may result in subtherapeutic concentrations of 2 ng/mL for some, and toxic concentrations of 30 ng/mL in others. To ensure tacrolimus concentrations stay within the therapeutic window, therapeutic drug monitoring is used. At UZ Leuven, newly transplanted patients undergo daily blood sampling in the morning. The central lab analyzes these samples through LC-MS/MS and reports tacrolimus concentrations to nephrologists by noon. During rounds, the dose is adapted accordingly.

Historically collected data from previous studies[136] was analyzed to quantify dose adaptation accuracy and precision. Based on N=315 patients, the fraction of patients in the target are summarized in Table **4** below. There is only a small number of patients in the target window. There is very few alarming under-dosing (below 5 ng/mL), but at least 1 in 10 patients is overdosed.

[135] Wallemacq et al., "Opportunities to Optimize Tacrolimus Therapy in Solid Organ Transplantation".

[136] Vanhove et al., "Pretransplant 4β-Hydroxycholesterol Does Not Predict Tacrolimus Exposure or Dose Requirements During the First Days After Kidney Transplantation".

***Table 4**: Overview of target attainment (in %) in N=315 patients transplanted at UZ Leuven between 2004 and 2014*

Day after transplant	<5	05-12	12-15	15-20	>20
1	8.5	38.2	16.3	14.7	22.2
2	0.6	21.5	16.7	28.9	32.2
3	0.3	23.0	22.4	33.9	20.4
4	0.3	31.6	29.4	29.7	9.0
5	0.3	40.0	33.2	22.3	4.2
6	0.0	47.1	34.4	15.6	2.9
7	0.3	54.0	31.8	11.3	2.6
8	1.3	56.2	31.4	11.0	0.0
9	1.0	59.1	30.9	8.6	0.3
10	1.1	61.8	28.1	8.2	0.7
11	0.0	64.0	27.6	7.5	0.8
12	0.0	63.8	27.1	9.0	0.0
13	0.0	60.7	29.5	7.7	2.2
14	0.6	58.3	30.7	9.8	0.6

By using a population pharmacokinetic model, concentration-time profiles can be predicted for individual patients. Based on only a few concentration samples, an accurate profile for the patient can be constructed, predicting future concentrations for a given regimen. This can then be used to find the ideal dosing regimen to hit a given target concentration at a future point in time.

Using retrospective data, such a population PK model was constructed for UZ Leuven patients. A dose adaptation software was implemented that uses this model. The software automatically receives patient data from the Electronic Patient Record system in use at UZ Leuven (Klinisch WerkStation / KWS) and can automatically adapt future prescribed doses.

It is our aim to investigate through simulations whether such a computer system can improve target attainment of tacrolimus concentrations in patients receiving de novo kidney transplantation.

8.3. Simulation objectives

To assess whether the use of a computer system to adapt daily tacrolimus doses…

1) *Increases speed of target attainment.* It is our hypothesis that a computer system would result in a faster attainment of target.
2) *Improves overall target attainment.* It is our hypothesis that a computer system would show increase the amount of trough samples in target overall.
3) *Improves average target attainment per patient.* It is our hypothesis that a computer system will quickly identify outlier patients and dose adapt them accordingly, while physicians adapt more slowly.
4) *Improves target attainment per day.* It is our hypothesis that a computer system will increase the number of patients in target for all days.

This evaluation will first be performed on the population level. To represent current standard of care, historic data will be used. To represent model-informed precision dosing results, an iterative cycle of fit-adapt-measure will be performed using the historic data.

In a second step, a random sample of 200 patients will be drawn from the population. This is repeated 10,000 times to obtain a sample of possible trial outcomes. These trial outcomes are all tested for significant results, which yields an estimate of statistical power.

8.4. Methods

For a detailed description of how the simulated concentration-time profiles were generated, we refer to Chapter 7.

8.4.1. Population analysis

630 patient profiles are available: 315 physician-adapted true concentration-time profiles, and 315 computer-adapted simulated concentration-time profiles. This is referred to as “the population”. Only 139 patients were observed for the full 14 days. For analyses where the time profile of a single

patient is important, this reduced subset was used. We assumed missing data was caused by non-informative dropout, so this does not introduce any bias in our results.

The difference between computer and physician dosing was analysed on three aspects: speed of target attainment per patient, average target attainment per patient, and overall target attainment per day. A graphical exploration was performed on the population first. We show the raw data as concentration-time profiles and as a scatter plot of concentration vs dose adaptation.

We then performed testing of the difference between physician and computer for the three aspects mentioned previously. Speed of target attainment per patient is shown as a Kaplan-Meier plot, and the p-value calculated using a Mantel-Haenszel test. Additionally, survival curves to reach two/three/four/… concentrations in target were plotted. Subsequently, we analysed ratio of days in target per patient. The normal, gamma, beta and logistic distributions were used to fit the data, and the most appropriate distribution was carried forward to quantify the difference between treatment arms. Finally, we show the proportion of patients in target per day. A proportions test was used to assert whether this distribution was different between arms.

As an alternative to dichotomizing concentrations into 'in target' vs 'out of target', squared distance to target window was used. To ensure findings can be translated to other target windows, data were first log-transformed. The empirical cumulative distribution function (ECDF) per day was then compared between physician and computer using a Kolmogorov-Smirnov test.

$$\text{if} \quad C < 12: \quad SquaredDistance = \big(log(C) - log(12)\big)^2$$

$$\text{if} \quad C \quad \text{in} \quad [12,15]: SquaredDistance = 0$$

$$\text{if} \quad C > 15: SquaredDistance = \big(log(C) - log(15)\big)^2$$

To avoid multiple testing, a generalized least squares model with within-patient correlation structure was used to fit the data. Different residual error variances were estimated for different days. To ensure normal data, data was first shifted by the mean (to avoid any '0' being present) and then box-cox transformed. A base model with only a fixed effect per day was fitted, and

subsequently amended with a single treatment effect, and with a treatment effect per day.

8.4.2. Statistical power of a clinical trial

To predict statistical power, the bootstrap method was used. 10,000 random trials were sampled by drawing 67 random patients from the physician-treated patients in the population, and 133 patients from the computer-treated patients in the population. Each trial was then analysed for difference between both arms, and p-value reported, using the following tests:

- TTE analysis using Mantel-Haenszel test
- Treatment effect on ratio of days in target per patient
- Proportion-test per day
- Significance of treatment effect in linear model on squared distance to target window

The resulting distribution of p-values is then graphically shown in an ECDF curve, representing study power to detect an effect at $p < X$.

8.5. Results

8.5.1. Graphical exploration

Looking at trough concentrations in Figure **31**, we observed a clear difference between physician and computer dosing. While physician-based dosing had median concentrations below the therapeutic window, computer-based dosing median concentrations were consistently in the therapeutic window, and close to the target of 13.5 ng/mL. The IQR was consistently smaller when using computer dosing.

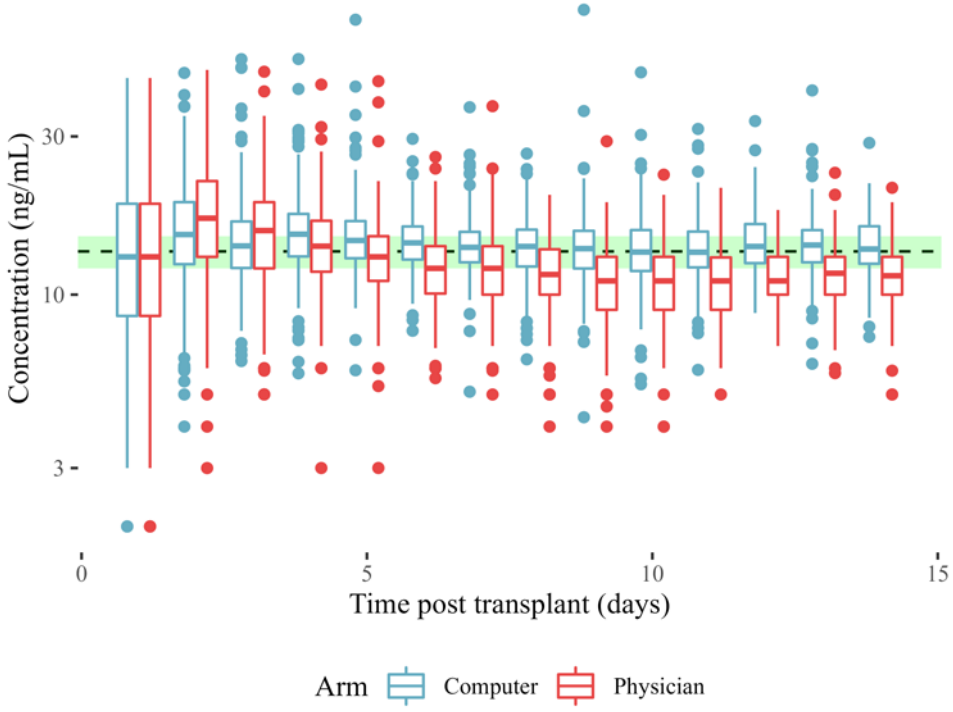


***Figure 31**: Trough concentrations with computer (red) or physician (blue) dosing. Therapeutic window shown in green.*

This can be explained by looking at dose adaptation in Figure **32**. Physicians generally underreact: concentrations between 5 and 12 ng/mL do not result in a high enough dose increase; concentrations between 15 and 20 ng/mL do not result in a high enough dose decrease. The computer seemingly overreacts: low concentrations result in high dose increases; high concentrations result in severe dose decreases. This is intended behaviour, as more pronounced dose adaptations are required to reach steady-state more quickly. The variability in computer dose adaptations for a given concentration shows the computer takes previous concentrations into account, as well as non-steady-state.

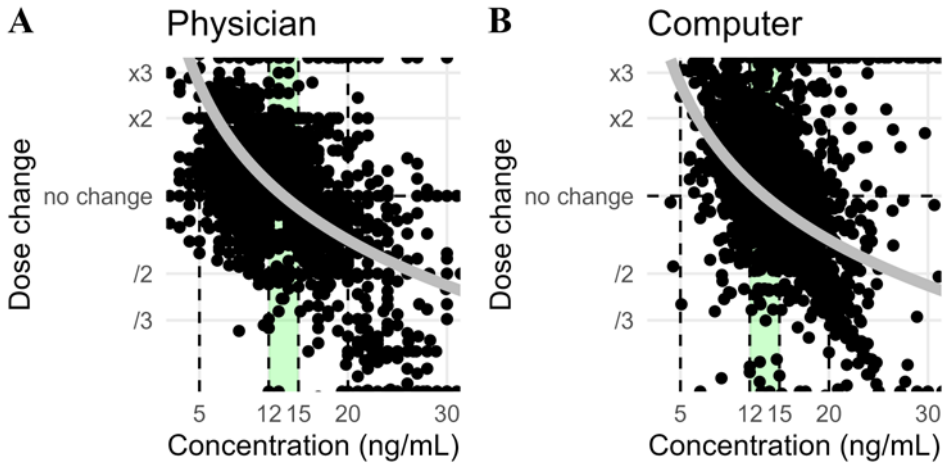


***Figure 32**: Physician dose adaptation (left) versus computer dose adaptation (right). The observed concentration (x-axis) results in a dose change on the subsequent day (y-axis). Assuming steady state, concentrations in target (green area, 12-15 ng/mL) should not induce a dose adaptation; lower concentrations require higher doses, higher concentrations require lower doses. The grey line shows the theoretical dose adaptation when following the rule of three.*

8.5.2. Time-to-event analysis

Looking at time-to-event in Figure **33** and Figure **34**, computer dosing significantly reduced time to reach target. After 1 dose adaptation, there is 8% more patients who reached target at least once. After 2 dose adaptations, this difference rises to 15%, a significant improvement over physician dosing. Median time-to-event decreased from day 4 to day 3.

This observation was maintained when evaluating time-to-2-samples-in-target, time-to-3-samples-in-target, etc. Computer dosing consistently reduced median event times: 1 day earlier for 1- and 2-samples-in-target, 3 days earlier for 3-samples-in-target. Less than 50% of patients had 4 or more samples in target with physician dosing, while computer dosing ended allowed more than 50% of patients to have 6 or more samples in target.

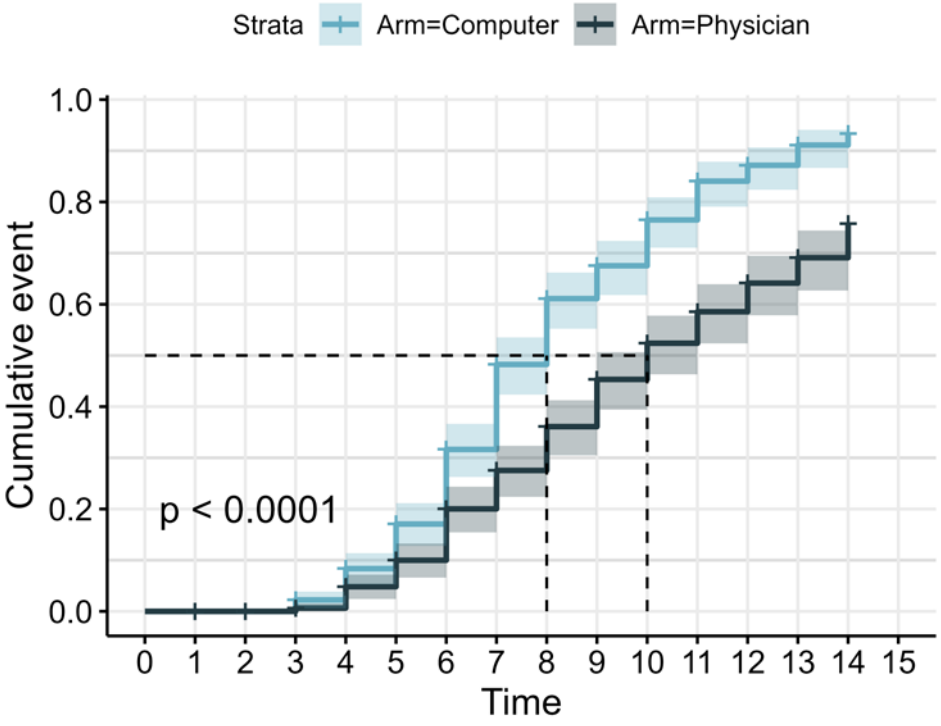


***Figure 33**: Time-to-event Kaplan-Meier plot showing how quickly a patient reaches at least 1 concentration in target*

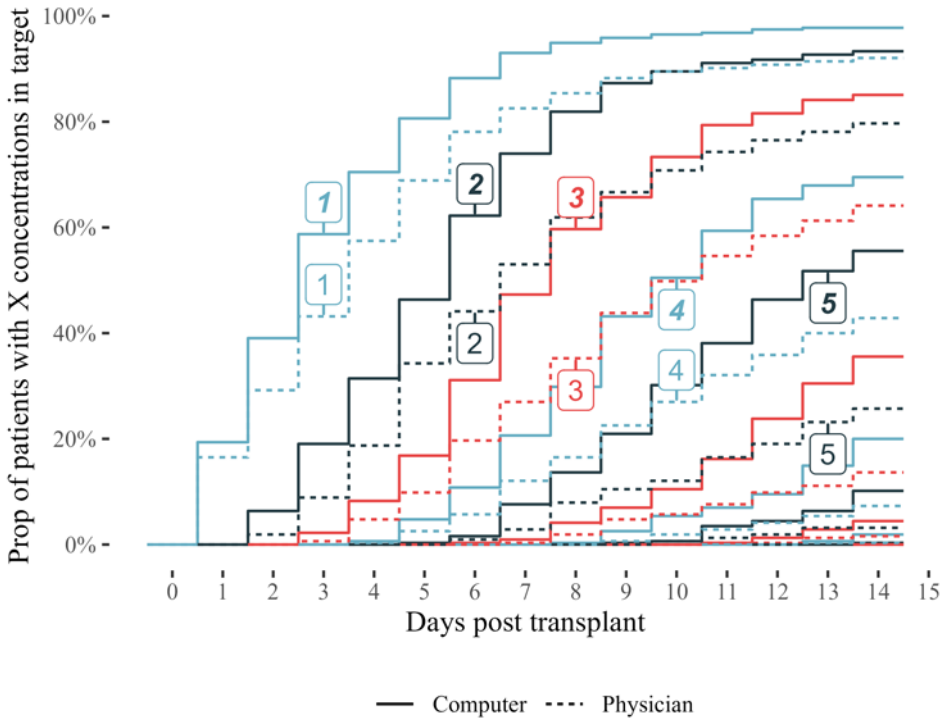


***Figure 34**: Time to event curves to reach 1, 2, 3, … concentrations in target. Computer (solid line, bold font) vs physician (dotted line, normal font).*

8.5.3. Time-to-event analysis: statistical power

Overall probability of success (Figure **35**) was predicted as 87.2% (at $p<0.05$) and 62.8% (at $p<0.01$) for time to reach 1 concentration in target. PoSS for time to reach 2, 3, 4, … concentrations in target was predicted even higher. PoSS decreased for time to reach 7, 8 or 9 concentrations in target, because few patients are predicted to reach this state.

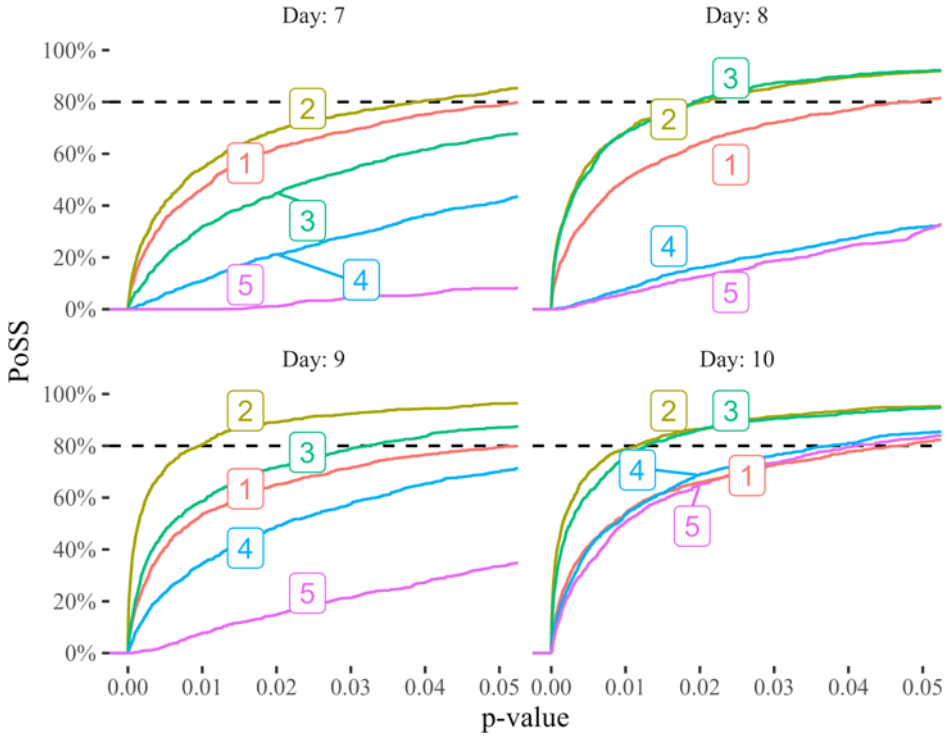


***Figure 35**: Power to detect a significant improvement in time to reach 1, 2, 3, … concentrations in target on day X.*

A cox proportional hazards model was also used to fit survival data. The model predicts different hazards depending on whether computer or physician dosing is used.

$$\lambda(t) = \lambda_0(t) \times e^{\beta * Physician}$$

e^{β} was estimated at 0.4742, meaning hazards are about half as low in the Physician arm. Schoenfeld residuals represent the difference between expected covariate distribution at a given time, and the actual covariate distribution. If the proportional-hazard model holds, these residuals should be constant over time: survival is as predicted by the model. Power did not improve however. The afore-mentioned Mantel-Haenszel test was still more robust.

8.5.4. Target attainment per patient

The computer increased the average ratio of days in target per patient (Figure **36**). The normal distribution fitted these data best, based on both Kolmogorov-Smirnov and Anderson-Darling statistics. We therefore used a t-test to compare both arms. The difference was highly significant.

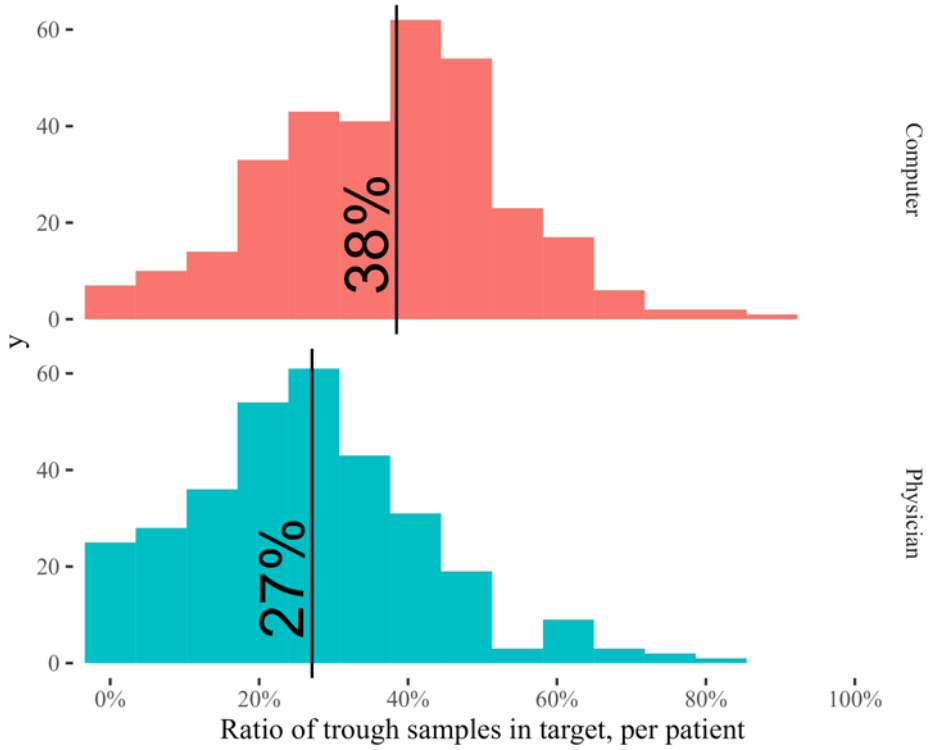


***Figure 36**: Ratio of trough samples in target, per patient. On average, 39% of samples are in target for the computer arm, versus 28% for the physician arm.*

8.5.5. Target attainment per patient: statistical power

All simulated trials showed a significant effect (Figure **37**). Expected observed effect size was +11.3% [+8%, +14.6% 95%CI].

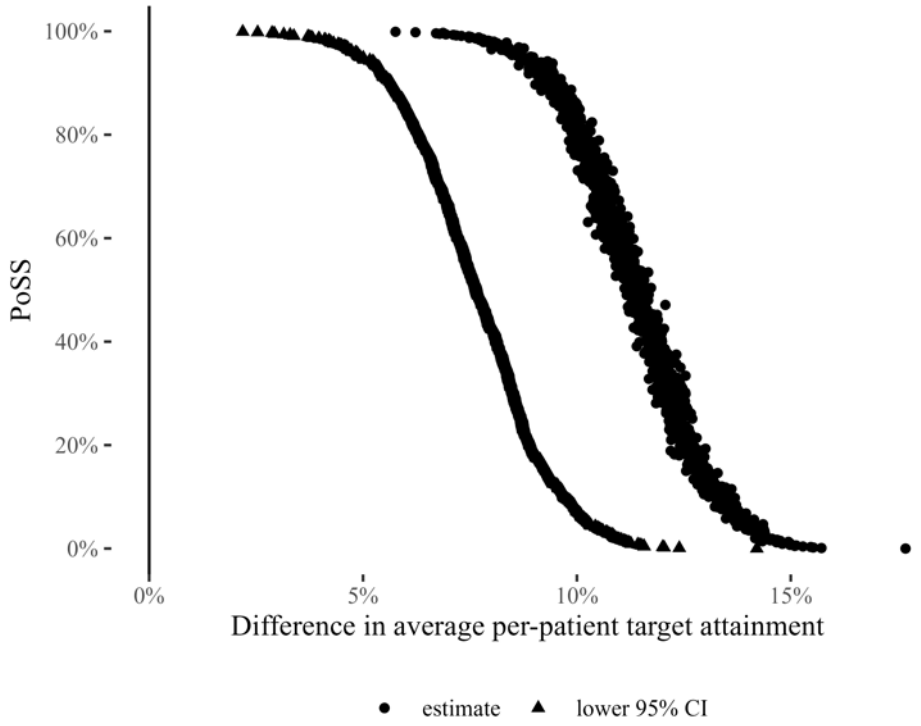


***Figure 37**: Estimate of difference in average per-patient target attainment for 1,000 random trials.*

8.5.6. Correlation between in target today and in target yesterday

There was a high correlation between being in target yesterday and being in target today. This correlation was highest in the physician arm. Indeed, there is a high probability of staying in target in the physician arm. Correlation is a false metric for performance. A good dose adaptation strategy ensures a high between-day correlation when in target, but a low between-day correlation when out of target.

As an approximation, a discrete markovian process was considered. Please note that this is a simplification, as this process can hardly be called markovian. Base transition probabilities in the population are shown in Table **5** below.

***Table 5**: Transition probabilities*

Physician	Today Out of target	Today In target
Yesterday Out of target	0.765	0.235
Yesterday In target	0.577	0.423
Computer		
Yesterday Out	0.605	0.395
Yesterday In	0.550	0.450

We fitted a discrete markov chain to each individual and plotted the estimated transition probabilities in Figure **38**.

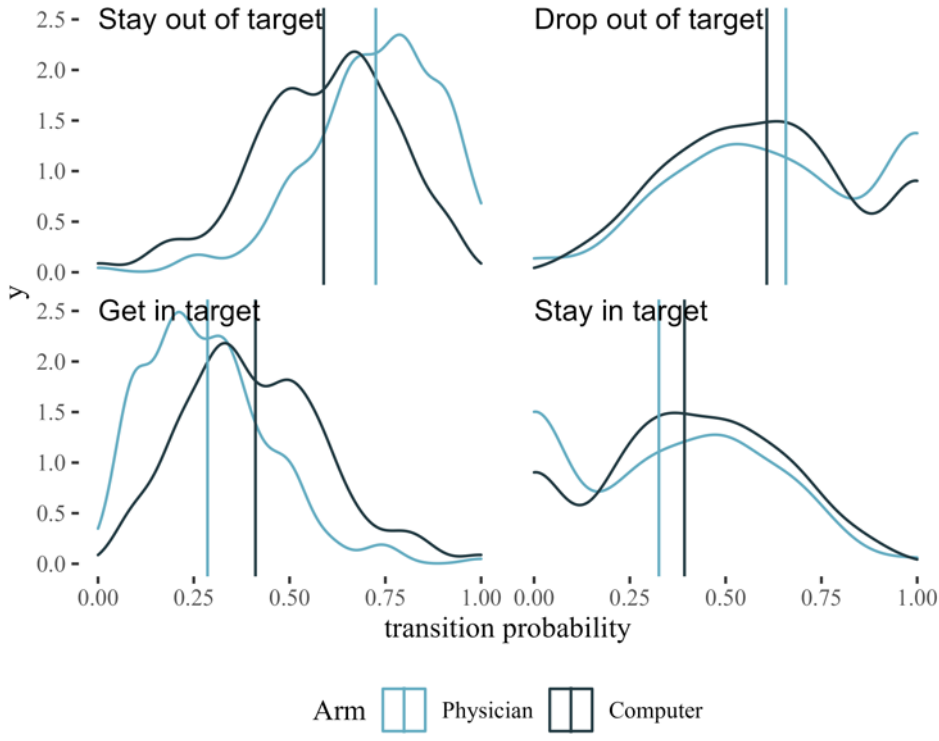


***Figure 38**: Density plots of individual transition probabilities, per arm*

The computer was predicted to only slightly improve the probability of staying in target (bottom right). The major change is in the bottom left of the figure: the computer increases the probability of getting into the target. Because the effect is only small, no power analysis was performed.

8.5.7. Proportion of trough samples in target

The amount of trough samples in target was analyzed and compared through a proportions test in Table **6**. There was a significant difference on days 2, 3, 6, 7, 8, 10, 11 and 12 when adjusting for multiple testing using Bonferroni correction ($p < 0.05 / 14 \rightarrow p < 3.5e\text{-}3$).

***Table 6**: Number of samples in target per day for Physician and Computer dosing. P-value to reject H0 (no effect) determined through t-test.*

Day	N	Physician	Computer	Increase	p.value
1	306	48	46	-0.7%	5e-01
2	311	48	84	+12%	3e-04
3	313	68	111	+14%	1e-04
4	310	95	110	+4.8%	1e-01
5	310	112	134	+7.1%	4e-02
6	308	113	149	+12%	2e-03
7	302	94	154	+20%	5e-07
8	299	89	129	+13%	5e-04
9	291	91	120	+10%	8e-03
10	267	70	109	+15%	2e-04
11	239	68	112	+18%	2e-05
12	210	55	89	+16%	3e-04
13	183	58	86	+15%	2e-03
14	163	47	71	+15%	4e-03

8.5.8. Proportion of trough samples in target: statistical power

When performing this test on a randomly sampled clinical trial however, there was almost never a significant difference (see Figure **39** below). Only for day 7 and 11 was a PoSS of more than 80% predicted at $p<0.05$.

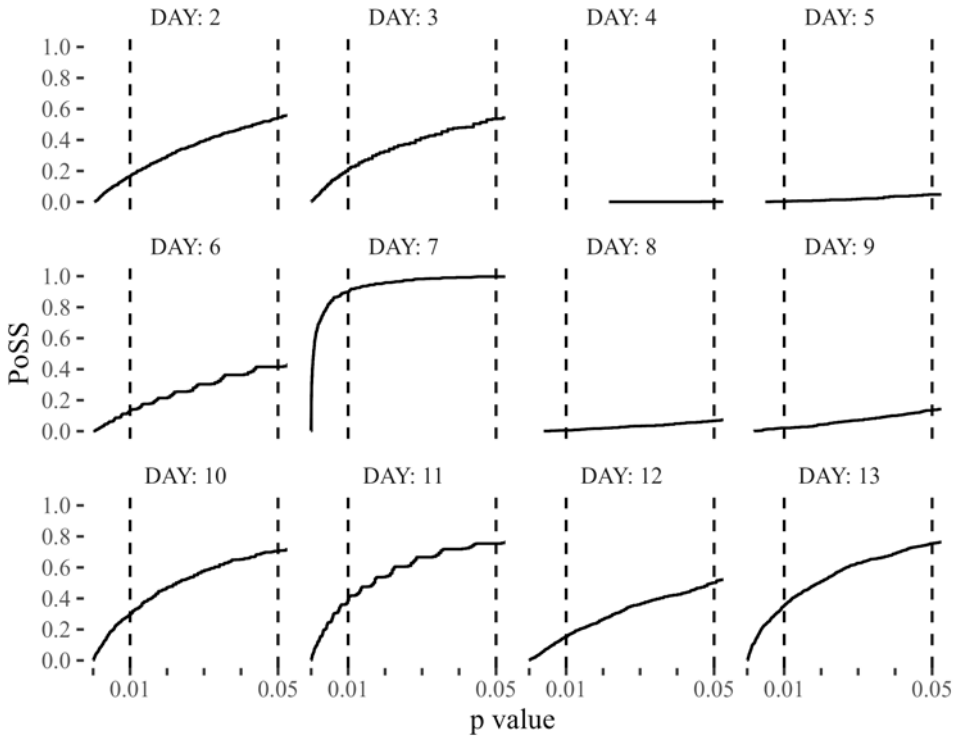


***Figure 39**: Power to detect a difference in proportion of trough samples in target per day.*

8.5.9. Squared distance to log-transformed target window and target attainment per day

When using the squared distance to log-transformed target window, significant differences for all days were observed (Figure **40** below). Statistical significance was determined by using a Kolmogorov-Smirnov test.

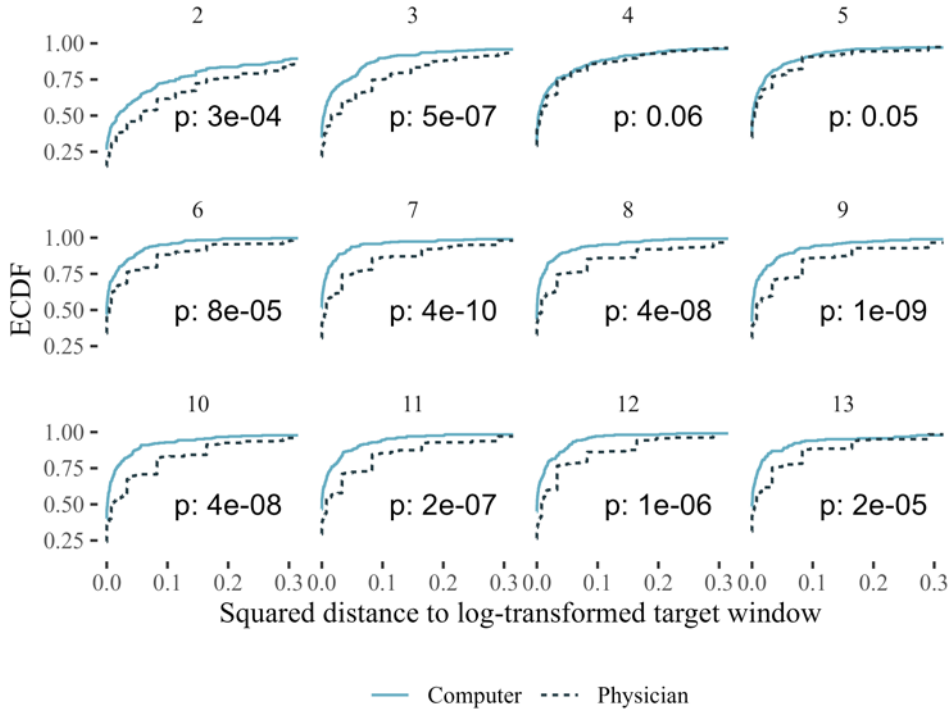


***Figure 40**: Difference between computer and physician in squared log-distance to target window, using KS-test*

8.5.10. Squared distance to log-transformed target window per day: statistical power

Statistical power for the KS test per day is shown in figure **41**. Only on day 3 and days 7 to 11 was a PoSS higher than 80% predicted for p<0.05.

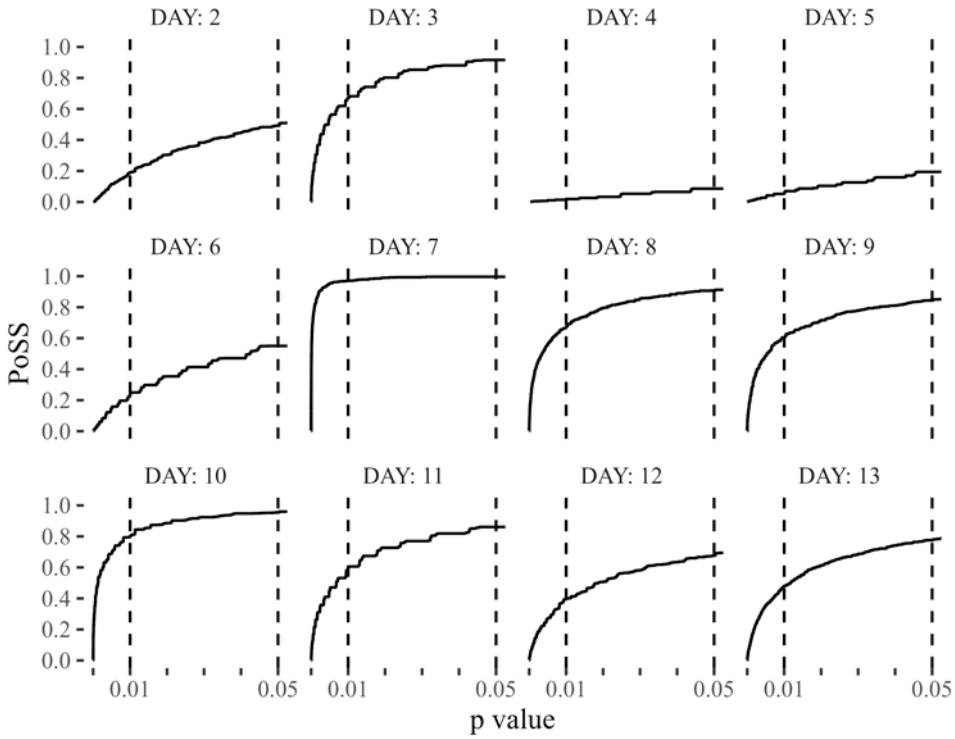


***Figure 41**: Power for KS-test of squared distance per day.*

8.5.11. Squared distance to log-transformed target window and linear model over all days

As expected, many concentrations were in the target window, and squared distance therefore showed a frequent number of '0'. This was shifted by the mean of the data to ensure a box-cox transform could be performed. The optimal lambda parameter was found as -1.86. Data was still far from normal, owing to a large amount of 0 in the original dataset (KS-statistic 0.22, AD-statistic 630).

A linear model was fitted to the transformed squared distance. The model building process is detailed below:

1) No clear algebraic relation could be found linking DAY to the outcome. DAY was therefore treated as a categorical variable.
2) Treatment effect was added as a single fixed effect.

3) Variance of squared distance is different between days. A variance structure allowing 14 different variances was used.
4) Squared distance is correlated between days for a given individual. A general correlation structure was used allowing for within-individual correlation between days.
5) A per-day treatment effect was added, as the difference between computer and physician clearly differs between days.

This resulted in the following model, shown below as R code:

```
gls(SquaredDeviation ~ Arm*DAY,
               # observations are correlated between days
               # within an individual, but uncorrelated
               # between different individuals
               corr = corSymm(form= ~ index | ID),
               # allow a different variance per day
               weights=varIdent(form = ~ 1 | DAY)
)
```

The model identified a significant treatment effect in the population (see Figure **42**). Lsmeans were calculated to predict both average and per-day treatment effect. This showed a significant treatment effect in the population, with per-day treatment effect significant at $p<0.05$, except for days 4, 5 and 13, with $p=0.15$, 0.26 and 0.078 respectively.

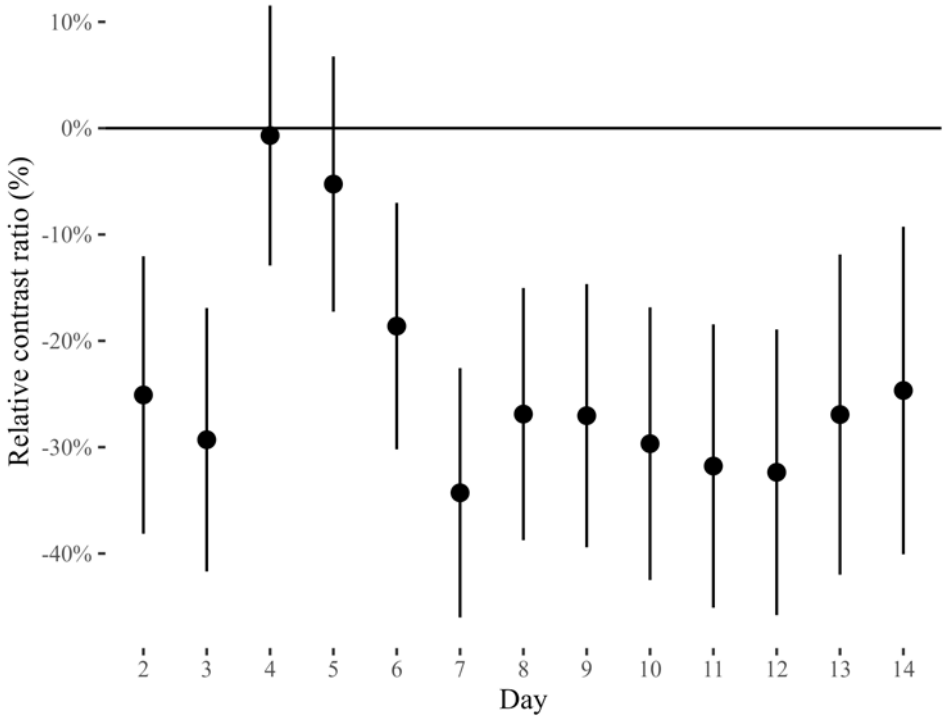


***Figure 42**: Effect size predicted by GLS in the population. Bars show 95% confidence interval.*

8.5.12. Squared distance to log-transformed target window: statistical power

Only 100 random trials were simulated, due to long runtimes. Statistical power to detect an average difference between arms was high: all trials showed p-values lower than 0.01. However, a per-day treatment effect was predicted to only reliably show at $p<0.05$ on days 6, 7, 8 and 11 (see Figure **43**).

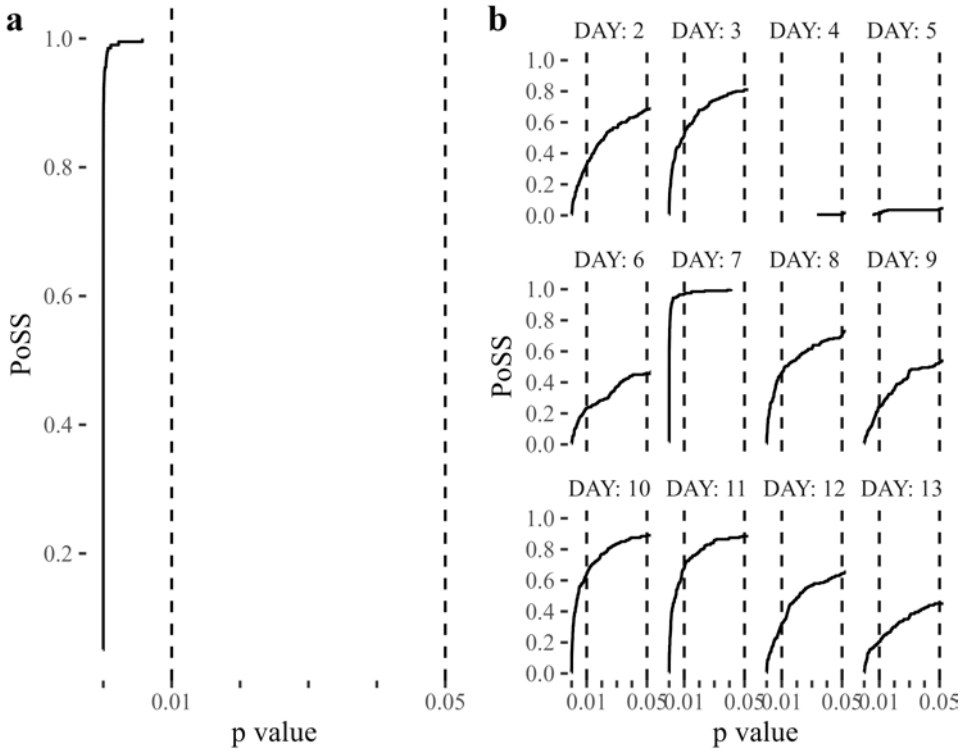


Figure 43*: Power to detect an improvement in overall (left) or per-day (right) squared distance to target window.*

8.6. Conclusion

In this study, we predicted how model-informed precision dosing improves tacrolimus target attainment in de novo kidney transplant patients in the first 14 days post-transplant. Specifically, the following four aspects were evaluated: time to target attainment, target attainment per patient, target attainment per day, and squared distance to target.

On a population level, model-informed precision dosing is predicted to improve all of these aspects significantly. After 1 dose adaptation, 8% more patients have reached the target. After a second dose adaptation, this increases to 15%. Per-patient target attainment increased from 28% to 39%. On all days, a higher proportion of samples is in target. Finally, a linear model predicts significantly lower squared distance from target. In conclusion, MIPD improves target attainment on all aspects.

Unfortunately, this does not mean we can show this effect in a study with a limited number of patients. An overall faster time-to-target can be reliably shown. This also applies to higher per-patient target attainment. Overall, an increase in target attainment can be shown, although this can only be reliably detected on day 6, 7, 8 and 11.

It should be noted that we used retrospective data as a representation of the control arm in a head-to-head study. Physicians may work more carefully than normal practice, resulting in increased target attainment. However, computer dose adaptation may likewise improve, thanks to improved data collection and reduced data error.

Our results are a blueprint for the study protocol of a randomized clinical trial comparing physician to MIPD dosing. This enables clinicians to take an informed decision about the clinical endpoint and expected results.

Chapter 9 Is tacrolimus precision dosing under azole co-medication worth it?

This chapter is based on the master thesis work of Mirthe Vincken, under daily guidance by Ruben Faelens and promoted by Pieter Annaert. R.F. wrote the text and supervised execution, M.V. executed the research, P.A. provided feedback and review of the thesis text. Dirk Kuypers provided the clinical and PK dataset and the proposal for the study.

9.1. Abstract

Introduction Tacrolimus is often co-administered with posaconazole or voriconazole anti-fungal agents, as immunosuppressed patients are particularly vulnerable for fungal infections. As both tacrolimus and -azoles are metabolized by CYP450, tacrolimus clearance decreases. This necessitates a tacrolimus dose adaptation. Unfortunately, such adaptation is not routinely performed nor recommended. Doctors rely on reactive TDM to eventually reach safe and effective tacrolimus concentrations. This study uses modeling of a retrospective dataset in lung and kidney transplant recipients to formulate proactive tacrolimus dose adaptation advice on start of -azole therapy.

Results A dataset of 122 patients was available: 91 lung transplant recipients and 31 kidney transplant recipients. A 2-compartment disposition with oral absorption described the tacrolimus concentration profiles well. Azole co-administration was estimated to reduce tacrolimus clearance 3-fold, with IIV of 55.5% CV% identified on this DDI effect.

Conclusion Based on these modeling results, we recommend that physicians proactively reduce the tacrolimus dose 3-fold when initiating -azole co-treatment. Due to the high remaining IIV on the DDI effect, reactive TDM is still recommended after -azole initiation to fine tune the dose.

9.2. Introduction

Tacrolimus is used to reduce risk of organ rejection in solid organ transplant recipients. This drug is a calcineurin inhibitor and thus inhibits a key step in the immune response, protecting the transplanted organ.[137] Tacrolimus therapy is not straightforward, as the wide inter-individual pharmacokinetic variability (IIV) and narrow therapeutic window requires regular patient follow-up in the form of concentration-based dose adaptation. Daily doses to reach effective and safe concentrations vary between 3mg and 30mg[138]. Further complicating this challenge are drug-drug interactions. Tacrolimus is

[137] Vincenti et al., "A Long-Term Comparison of Tacrolimus (Fk506) and Cyclosporine in Kidney Transplantation".

[138] Staatz and Tett, "Clinical Pharmacokinetics and Pharmacodynamics of Tacrolimus in Solid Organ Transplantation".

metabolized by CYP450 enzymes, primarily CYP3A5 and CYP3A4. Even if TDM manages to adapt the dose to varying inter-patient CYP450 activity, concomitant drug influencing these enzymes -either as inhibitor/substrate or inducer- may result in unsafe or ineffective tacrolimus concentrations[139].

One such drug class is azole antifungal agents, such as voriconazole and posaconazole. Fungal infections are common in immunosuppressed patients.[140] Invasive aspergillosis, the most common infection-causing pathogen, occurs in as much as 15% of solid organ transplants, and mortality rates are approx 22%. In lung transplant, mortality rises as high as 67-82%. As such, antifungal agents are essential as prophylaxis, with even higher doses required for treatment. Most azoles are notorious for their potential to inhibit CYP3A4. Therefore, they pose a major drug-drug interaction with tacrolimus, with dose-normalized concentration observed to increase 4-fold with voriconazole co-therapy, and 3-fold with posaconazole co-therapy[141]. To correct for this, tacrolimus doses should be proactively reduced.

However, as reported by Vanhove et al,[142] proactive dose adaptation is not commonly performed, nor is it performed in the same manner in all patients. Although the optimal individual dose is eventually identified through concentration-based dose adaptation, an optimal initial dose adaptation would ensure improved tacrolimus safety during the first days post azole initiation.

Vanhove et al also showed high variability in the effect size of drug-drug interaction. Mean dose-normalized concentration ratio with versus without azole co-therapy ranged from 186% to 796% (5th percentile and 95th percentile). This beckons the question whether this variability may be predicted, either through covariates or through correlation with concentration measurements pre-initiation of azole. The presented work

---

[139] Zhang et al., "Effect of Voriconazole and Other Azole Antifungal Agents on Cyp3a Activity and Metabolism of Tacrolimus in Human Liver Microsomes".

[140] Kabir, Maertens, and Kuypers, "Fungal Infections in Solid Organ Transplantation".

[141] Vanhove et al., "Determinants of the Magnitude of Interaction Between Tacrolimus and Voriconazole/Posaconazole in Solid Organ Recipients".

[142] Vanhove et al.

uses population PK (popPK) non-linear mixed effects modeling on a retrospective dataset to answer these questions and identify the optimal dose adaptation strategy.

9.3. Methods

9.3.1. Source dataset

A retrospective dataset collected by Vanhove et al (2017) was used. Collected patient data ranged from 2007 to 2015. This multicenter study consisted of both lung and kidney organ transplant recipients. Azoles were administered both as prophylaxis early post-transplant (<90 days), or as treatment for a suspected fungal infection. For more details, we refer to the publication by Vanhove et al.

Tacrolimus concentrations were observed at the following timepoints: at least three timepoints prior to azole initiation, all available timepoints during azole treatment, and 7, 14, 30 and 90 days post azole discontinuation.

9.3.2. Modeling

A 2-compartment disposition with oral absorption was used to describe the data. As only sparse data was available, several parameters were fixed to values from literature.[143] Absorption rate K_a, intercompartmental clearance Q and peripheral comparment volume $V2$ were fixed at $0.53h^{-1}$, $70L/h$ and $1300L$ respectively. Clearance CL and central compartment volume $V1$ were estimated.

Presence of antifungals was modified as an ON/OFF modifier on clearance using the following formula, with $\beta_{azole,CL}$ the covariate effect and $AzolePresent$ at value 0, 1, and 0 before, during and after azole treatment respectively:

$$CL = CL_{base} \times e^{\beta_{azole,CL} \times AzolePresent}$$

[143] Musuamba et al., "A Simultaneous D-Optimal Designed Study for Population Pharmacokinetic Analyses of Mycophenolic Acid and Tacrolimus Early After Renal Transplantation".

Inter-individual variability was estimated on CL and $V1$, with correlation. Inter-individual variability was also estimated on $\beta_{azole,CL}$. A combined additive and proportional residual error model was used.

Model fit was evaluated through prediction-corrected VPC, as well as standard GoF metrics such as longitudinal individual weighted residual plots and log-likelihood. A stepwise covariate search was performed on a subset of clinically relevant covariates. Covariates leading to a statistically significant model were included.

9.3.3. Simulation

To characterize DDI-related dose adaptation performance, we focused on $\beta_{azole,CL}$. This allowed us to ignore all tacrolimus PK variability unrelated to DDI. The individual posthoc estimates CL_{ref} and $V1_{ref}$ from the final popPK model on all available data were considered as reference values. As prediction strategy, standard of care (SoC), model-informed fixed effect, model-informed covariate-based effect and model-informed concentration-based effect were evaluated. The mean squared relative error was compared for each adaptation strategy.

The relative error in clearance prediction post-azole initiation can be directly related to error in steady-state concentration. If we assume a trough concentration of 10 ng/mL pre-initiation, the concentration post-initiation C_{post} is as follows:

$$C_{post} = 10 \times \frac{CL_{pred}}{CL_{ref}}$$

A reclassification table was created to show how SoC compared to new dose adaptation strategies. Patients were divided in 5 categories: no efficacy (<5ng/mL), at risk for efficacy, in target (8-12ng/mL), at risk for toxicity, and toxicity (>15ng/mL).

9.3.4. Software

All statistical analysis was performed using R 4.0.4. Non-linear mixed effects modeling was performed using Monolix 2019.

9.4. Results

9.4.1. Source dataset

126 patients were available for inclusion. Exploratory data analysis identified four patients who received azole therapy prior to transplantation; these were excluded. Two implausible measurements were also excluded.

The final dataset consisted of 122 patients: 91 lung transplant recipients, with 25 patients suffering from cystic fibrosis (CF), and 31 kidney transplant recipients. 24 lung transplant patients received posaconazole, all other patients received voriconazole.

Tacrolimus trough concentrations (regular and dose-normalized) are presented in figure **44**. The figure shows insufficient dose adaptation post-azole initiation, with tacrolimus concentrations rising above 15 ng/mL for a quarter of patients. Only after a considerable period are doses adapted to reach safe concentrations. Underexposure is apparent at azole discontinuation, as the dose is not proactively adapted to correct for renewed tacrolimus clearance. Dose-normalized concentrations clearly show the reduced tacrolimus clearance during comedication with azoles. As doses were continuously adapted, dose-normalized concentration is not steady state. Therefore, it is not a perfect surrogate for clearance, explaining the high variability between days.

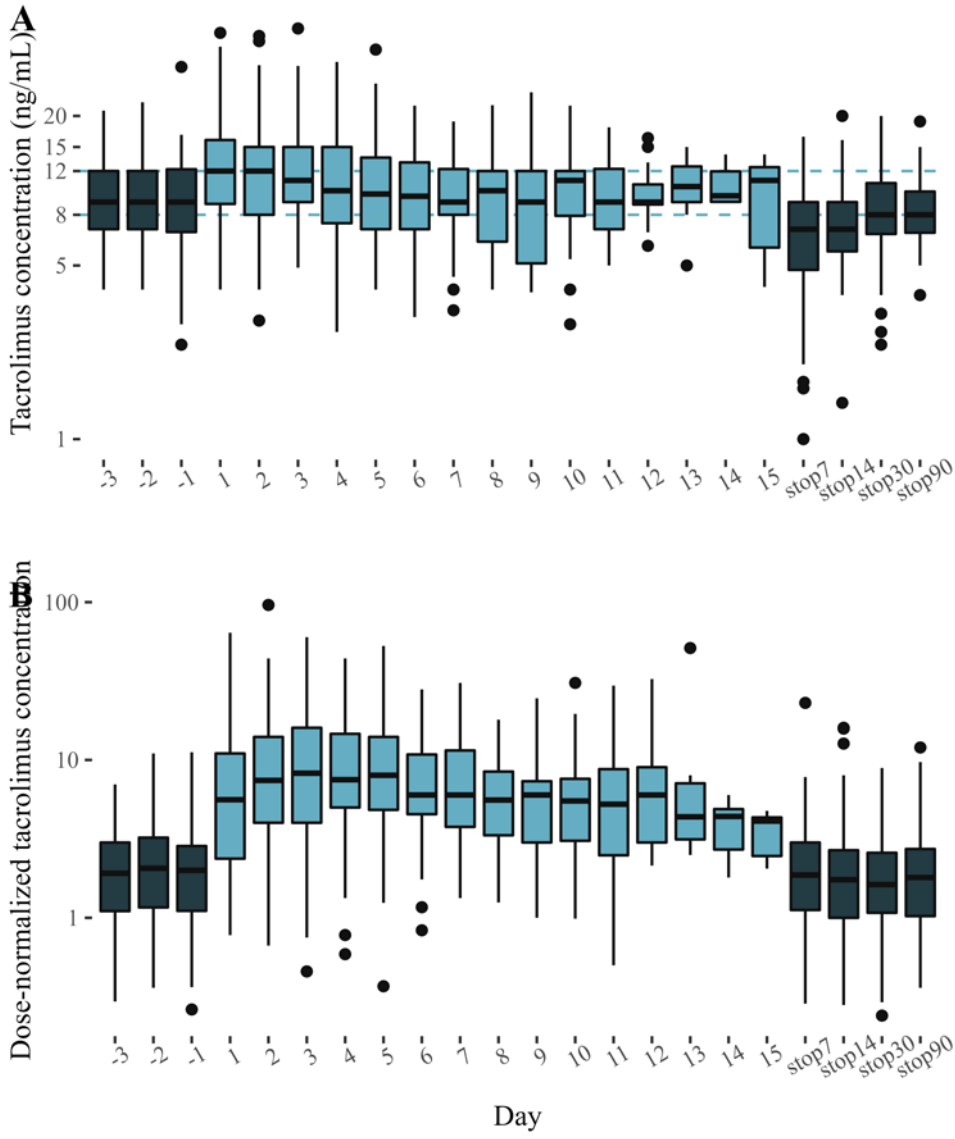


Figure 44*: Tacrolimus concentration (a) and dose-normalized concentration (b) before (dark blue), during (light blue) and after (dark blue) azole therapy. Boxplot shows median (center line), 25th and 75th percentiles (edges of box), data range (whiskers). Outliers further than 1.5 times inter-quartile range are plotted as points.*

9.4.2. Modeling

The base model had acceptable fit with low residual errors. Clearance and central compartment volume were estimated at $20.6 L/h$ and $2820 L$ respectively. The azole effect on clearance was estimated at 0.319, meaning clearance was at 31.9% of base clearance during azole comedication. IIV was estimated high, at 98.6% and 192% for clearance and central compartment volume respectively. Parameter estimates are summarized in Table **7**.

This model was further improved by inclusion of CYP3A5 effect on clearance (18.8 L/h for expressors, 13.2 L/h for non-expressors), IIV on the azole effect (55.5%), and correlation between IIV of V1 and CL (-0.442). This reduced 2log-likelihood by 70.4 points.

No correlation could be found between tacrolimus disposition IIV and azole effect size IIV. There was no statistical evidence for an azole effect size linked to azole concentrations. Further covariate search focused only on azole effect IIV. The following covariates were tried: age, proton-pump inhibitor comedication, bodyweight, CYP2C19*2 genotype, CYP2C19*17 genotype, CYP3A5*3 genotype and transplantation type. Only PPI resulted in a statistically significant improvement.

***Table 7**: Parameter estimates of the base and final population pharmacokinetic model.*

Parameter	Base model (OFV=8636.12)	RSE	Final model (OFV=8565.72)	RSE
Typical values				
Ka [/h]	0.53	fix	0.53	fix
CL [L/h]	20.6	9.4%	18.8	9.3%
V [L]	2820	25%	988	34%
Q [L/h]	70	fix	70	fix
V2 [L]	1300	fix	1300	fix
Covariate effects				
Azole on CL	0.319	0.81%	0.325	6.2%
CYP3A5 on CL			0.703	36%
Inter-individual variability				
CL	168%	6.9%	146%	7.4%
V1	585%	11%	1150%	10%
Azole on CL			55.5%	11%
corr V1,CL			-0.442	22%
Residual error				
Additive [L/h]	0.284	7%	0.261	7.8%
Proportional	0.138	5.9%	0.134	5.9%

Inter-individual CV% was calculated as sqrt(exp(omega^2)-1). Relative standard error (RSE) was determined through importance resampling.

9.4.3. Simulation

Based on the retrospective dataset, standard of care dose adaptation was determined to be a 50% dose reduction. This resulted in a MSE of 0.468. The model identified a typical $\beta_{azole,CL}$ of 0.319. For ease of clinical implementation, this was rounded to 0.33 (divide dose by 3). This resulted in a far lower MSE of 0.099, indicating it is a superior dose adaptation. A dose reduction dependent on PPI comedication (0.318 in presence of PPI, 0.408 without) further reduced MSE to 0.0734. More details are available in Table **8**. Relating this to clinical impact, Figure **45** shows clear superior performance for dividing the dose by 3.

Table 8*: Calculated mean squared error for different prediction models. P-value shown for a t-test proving difference of means with the base model in bold-font.*

	MSE	p.value
PPI covariate effect	0.047	0.019
CYP2C19_17 covariate effect	0.060	0.155
PPI and Age covariate effect	0.064	0.306
CYP3A5 covariate effect	0.069	0.513
Age covariate effect	0.070	0.535
BW covariate effect	0.075	0.786
Tx covariate effect	0.079	0.985
No covariate effect	**0.080**	**NA**
CYP2C19_2 covariate effect	0.080	0.963

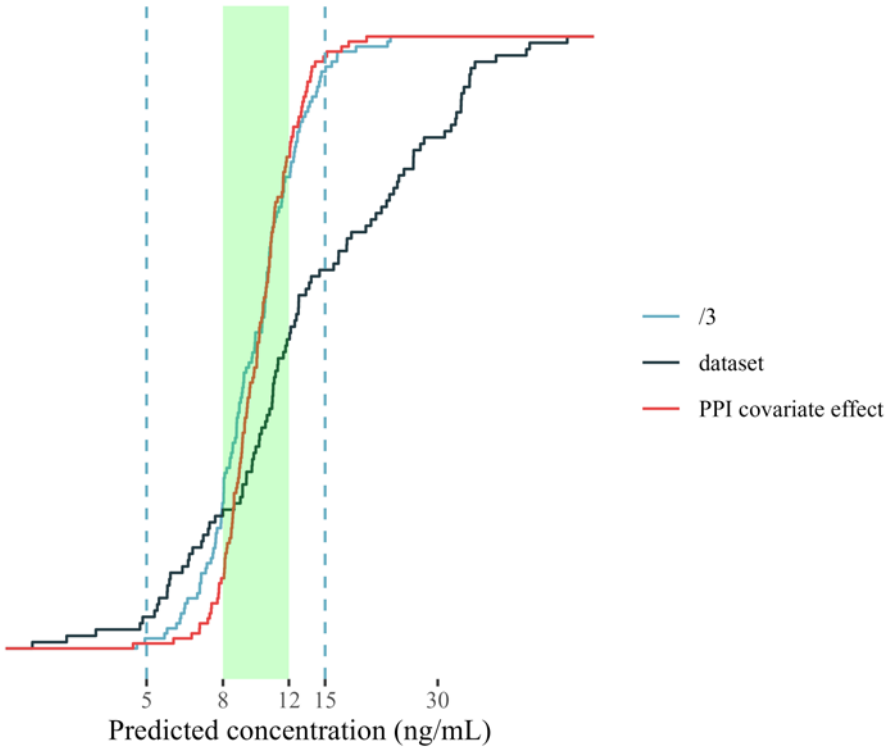


***Figure 45**: Outcome comparison between dosing strategies, as simulated trough concentration post-initiation. Lines show empirical cumulative distribution function for the recommended strategy, the currently performed dose adaptations, and the optimal PPI covariate model. Target window in green, edges of toxicity and efficacy as vertical dotted lines.*

9.5. Discussion

This work establishes a clear recommendation for tacrolimus dose adaptation post-azole initiation. The tacrolimus dose should largely be divided by 3. More precisely, the dose should be adapted to 31.8% with PPI co-medication, or to 40.8% without PPI.

The identified pharmacometric model is in line with literature. Parameter estimates for tacrolimus disposition broadly match tacrolimus models for

kidney and lung transplant.[144] It should be noted that cystic fibrosis (CF) tends to cause malabsorption, and thus results in a decreased tacrolimus bio-availability (47% slower absorption and 63% less bio-availability).[145] CF did not seem a significant covariate in this dataset, which is partly explained by the lack of rich data, and partly by the presence of pancreatic enzyme replacement therapy mitigating this effect.

A very large inter-individual variability of 1150% was identified on V1, with typical value of V1 dropping about 3-fold between base and final model. All patients in this dataset are at tacrolimus steady-state. Steady-state trough concentrations yield little information to accurately estimate distribution volumes. This is also reflected in the high IIV correlation between $V1$ and CL of -0.442.

The identified optimal dose adaptation matches those by Chheda, Tarleton, and Eidem[146], who studied tacrolimus and voriconazole in heart transplant recipients. They reported an average 180% increase in dose-normalized concentrations in the voriconazole arm, resulting in 67% lower dose requirement.

It should be noted that our physicians adapt doses widely. We did not find any correlation with patient covariates to explain this variation. A histogram of dose adaptation is available in Supplementary Materials. We hypothesize some of the variability is caused by ongoing dose adaptations early post-transplant as fungal infection prophylaxis is initiated in parallel. A high number (20%) of patients received no dose adaptation whatsoever. We strongly recommend the implementation of a straightforward and consistent dose adaptation strategy "divide by 3" rather than more complex systems, to maximize chances of adoption by physicians.

---

[144] Brooks et al., "Population Pharmacokinetic Modelling and Bayesian Estimation of Tacrolimus Exposure".

[145] Monchaud et al., "Population Pharmacokinetic Modelling and Design of a Bayesian Estimator for Therapeutic Drug Monitoring of Tacrolimus in Lung Transplantation".

[146] "Targeted Aspergillus Prophylaxis With Voriconazole in Heart Transplant Patients".

We also found that IIV on the azole effect remained high. Therefore, reactive concentration-based dose adaptation remains necessary. Remarkably, when average daily dose with azole comedication was compared to the dose pre-azole, an average of 0.3574 was found. This closely matches the model-predicted azole effect, and indeed shows that concentration-based dose adaptation also refines the dose appropriately, albeit slower than our suggested strategy.

Whether this dose recommendation may affect clinically meaningful improvements is an open discussion. The relevance of PK targets and their translation to clinical outcomes is still an area of active research. Miano et al.[147] demonstrated high tacrolimus exposure is associated with higher acute kidney injury risk, yet could not show increased cellular rejection risk at low exposures. More research is needed to elucidate tacrolimus exposure-response in lung transplant. Furthermore, we have not studied azole PK in this work. A reduction in tacrolimus dose may also impact azole disposition, resulting in increased availability of CYP enzymes, higher azole clearance and lower exposure.

Finally, this work focused on azole initiation. Perhaps even more important is the azole discontinuation, as patients are usually sent home with fewer opportunities for tacrolimus concentration-based dose adaptation. We did not show any correlation between azole concentration and magnitude of DDI effect, as these were clinically relevant concentrations at or above efficacy. During washout, a concentration-dependent DDI may appear. Azoles bind strongly to P450 enzymes, blocking the catalytic cycle (suicide inhibition).[148] The DDI we observed may be purely dependent on binding rates of azole to P450, and endogenous production of P450 enzymes. At lower concentrations, competitive binding may be present instead. An estimate of azole washout, azole concentration-dependent DDI effect, and associated tacrolimus dose increase scheme should prevent tacrolimus underexposure for recently transplanted patients.

In conclusion, this work presents clear evidence for an increased tacrolimus dose reduction “divide by 3” when initiating azole treatment. We urge

---

147 “Early Tacrolimus Concentrations After Lung Transplant Are Predicted by Combined Clinical and Genetic Factors and Associated With Acute Kidney Injury”.

148 Balding et al., “How Do Azoles Inhibit Cytochrome P450 Enzymes?”

transplant doctors to implement this in practice, and evaluate its impact on tacrolimus target attainment and clinical outcomes.

9.6. Supplementary Materials: dosing histogram

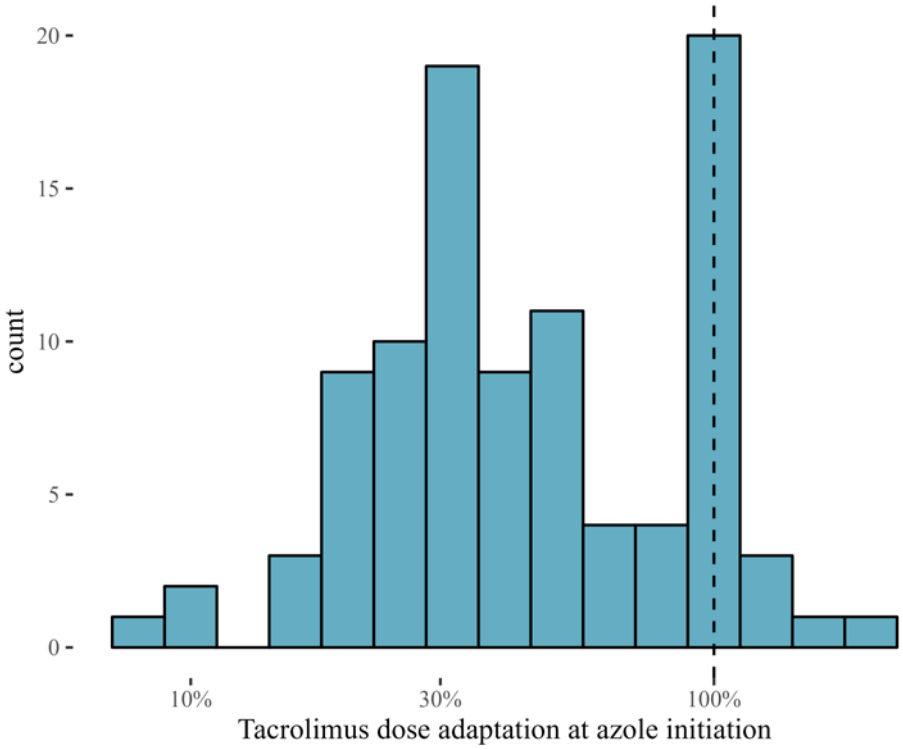


***Figure 46**: Tacrolimus dose adaptation post azole initiation as performed by physicians from 2007 to 2015. In 19 of 97 patients, no proactive dose adaptation is performed.*

Chapter 10 Discussion

In this chapter, we evaluate whether the overarching goals described in Chapter 1 were achieved. We summarize strengths and weaknesses of the employed approach, and formulate recommendations and opportunities for future research.

10.1. Evaluation of objectives

10.1.1. Simulate MIPD: general framework

Objective 1 – Achieved

Develop software and methodology to quantify model predictive performance, simulate outcomes applying MIPD to a virtual patient population, and simulate clinical trials including MIPD

A flexible framework was developed to support development of precision dosing during model development, population simulation, clinical trial design, and software implementation. The framework was sufficiently generic to allow application in different compounds (tacrolimus and infliximab) and different dosing targets (trough concentration target window for tacrolimus, optimal probability of endoscopic improvement for infliximab). The open-source nature allows developing previously unexplored functionality, such as dosing interval adaptation.

10.1.2. Simulate MIPD for infliximab induction therapy in ulcerative colitis patients

Objective 2 – Achieved

By performing an in silico population simulation of infliximab precision dosing, the improvement on PD outcomes can be quantified.

We have presented a structured roadmap for precision dosing development, accompanied by use cases demonstrating the value of each intermediary step. Early identification of MIPD potential in model development was applied for proactive dose adaptation of tacrolimus at initiation of posaconazole or voriconazole treatment in Chapter 1. The model identified an optimal mean proactive tacrolimus dose adaptation, but IIV of this effect

could not be linked to any measure before initiation, thereby precluding model-based proactive dose adaptation. Model refinement through predictive performance evaluation showed that complex models for tacrolimus PK did not outperform simpler models using MPC/MIPD (Chapter 7). Population simulation of precision dosing showed precision dosing of infliximab induction therapy fell short of expectations (Chapter 1). No detectable[149] benefit was predicted, while costs increased. Finally, simulation of clinical trial design identified the risk of an inconclusive trial on tacrolimus MIPD, thereby providing evidence to double the enrollment to 200 patients, improving predicted probability of study success.

10.1.3. Simulate MIPD for tacrolimus in de novo kidney transplant recipients early post-transplant

Objective 3 – Partial success

By quantifying the improvement MIPD brings to tacrolimus target attainment, the clinical benefit vs required effort can be evaluated. A targeted clinical trial to show this clinical benefit can be efficiently designed.

We have demonstrated the capability to relate precision dosing directly to clinical benefit in the infliximab usecase (under assumptions), and further provided a theoretical example optimizing for individual clinical utility in Chapter 2.

For tacrolimus, while a clear impact on surrogate endpoints was demonstrated, no impact on clinical benefit could be shown. The correlation between exposure and clinical benefit remains poorly defined for this therapeutic area. To this day, different hospitals use different target ranges. While current retrospective studies try to relate widely varying PK to differing clinical outcomes, this discounts individual variability in PD, which may explain why no universal PK target could yet be identified. A concentration-controlled trial could serve to characterize PD variability and unravel the seemingly variable exposure-response relationship of tacrolimus. We clearly demonstrated this type of trial is feasible through the use of MIPD.

[149] For a dichotomous YES/NO endpoint, only the mean outcome can be observed in a clinical trial: “X% of patients were cured”. Mean outcomes did not improve with infliximab MIPD.

10.1.4. Build a tacrolimus MIPD software tool

Objective 4 – Achieved

If MIPD is predicted to deliver a clinically significant benefit to patients, a software tool will bring this technology in the hands of physicians.

It was surprisingly easy to leverage the software framework to build a precision dosing tool. Practically implementing this tool allowed us to discover two surprising pitfalls. Any software tool will always require coordination with the clinical workflow; needing change at either (or both) ends. We have demonstrated an open-source framework enables adapting the software tool itself, minimally disrupting established clinical practices. Second, developing a medical device was met with severe legal concerns. Before 2017, medical device software was largely exempt from regulations as long as the final dosing decision rested with a human physician. With EU regulations 2017/745 (Medical Devices) and 2017/746 (In-Vitro Diagnostics) entering into vigor, software medical devices are severely restricted, rendering open-source development all but impossible.

10.1.5. Transpose this approach to other compounds

Objective 5 – Failed to demonstrate

The design of the software and approach should allow easy extension to other compounds.

The software was designed to be flexible and easily allow application to other compounds. This flexibility was demonstrated by implementing two usecases: tacrolimus dose adaptation for kidney transplant recipients, and infliximab dose adaptation for ulcerative colitis induction therapy.

However, these usecases were implemented by the same group who developed the framework. The inherent flexibility would be much more convincing in a collaboration with external parties. Unfortunately, custom license agreements were needed to share the full framework with other parties. Although there was demonstrated interest[150], parties failed to follow

[150] An industry pharmacometrics group wanted to use the user interface to educate the clinical team in Phase I on dose escalation. A UK-based academic

through, hindered by the additional legal burden to set up a custom license agreement. Collaboration was made impractical by legal risks, hurting our ability to demonstrate software flexibility.

Even within the University of Leuven, adoption was difficult, despite organizing workshops and tutorials. We identified three potential reasons. Without a large userbase supporting the open-source version, maintenance is not guaranteed. Second, the use of opensource software was deemed a risk for potential future commercialisation of MIPD software, even when many counterexamples exist[151]. Finally, the "Not Invented Here" syndrome may hurt adoption by research groups, who -as identified by-[152] may be reluctant to share process knowledge, may prefer internal knowledge rather than dependence on external factors, may be reluctant to set up external collaborations, or may feel threatened by competition. Should all stakeholders be willing, we are certain the full publication of our successful proof of concept software on tacrolimus -in active use at the hospital- could nevertheless make a dent in this domain, serving as a blueprint easily adapted by others for different compounds.

Notwithstanding the above, further development of the tacrolimus precision dosing software is planned, extending into other solid organ transplants. This demonstrates incorporation of other models, even if no extension to other compounds is planned as of yet.

10.2. Model building

We have shown throughout this work that model building for precision dosing shows distinct differences from population modeling. As theorized in section 4.1, the predictive power of a model is key. This is apparent in each presented case study. For tacrolimus PK, a hematocrit-standardized model

group wanted to use the software to build MIPD software for the pediatric ward of a hospital. A New-Zealand-based group expressed interest to review the software, and use this to improve their own software product.

[151] Nagy, Yassin, and Bhattacherjee, "Organizational Adoption of Open Source Software".

[152] Grosse Kathoefer and Leker, "Knowledge Transfer in Academia".

showed a better fit[153], but introduced bias if individual future hematocrit was not jointly modeled and predicted. Even then, a simpler model using MPC/MIPD showed superior predictive performance, even when population fit was worse. An already established model may need to be re-evaluated for use in MIPD. To predict MIPD performance for infliximab, exposure-associated covariates that could not be jointly modeled were removed, increasing IIV. Future studies may focus on jointly modeling continuous PD endpoints, thereby increasing MIPD accuracy and potentially reversing our findings.

We proposed to investigate autocorrelation as an extra GoF diagnostic. This allowed us to identify an opportunity for improvement in tacrolimus PK, developing MPC/MIPD as a pragmatic solution. The method was defined through pragmatism, and can certainly be refined in future research. Specifically, the tuning parameter for the inter-individual variability in future occasions should be further investigated. Different strategies should be compared, using either (a) the a priori IIV ω, (b) the parameter imprecision ϕ, or (c) a combination of both, possibly influenced using identified auto-correlation.

As shown by McDougall et al.[154], model misspecification will generally not impact MIPD performance. While a pharmacometrician focuses on improving population fit, predictive performance -what actually matters for MIPD- falls by the wayside. A diagnostic toolset to investigate predictive performance, and a modeling process that leads to high predictive performance, is urgently needed.

In targeting optimal target attainment, or striving for optimal target concentration,[155] we fail to address the elephant in the room. PK is only an intermediary step, and we should target optimal clinical utility instead. Unfortunately, the exposure-response relationship was not fully characterized during drug development for many compounds on the market to date. The litmus test for drugs to gain market access should be more strict:

---

[153] As defined by an improved log-likelihood, and a better visual predictive check.

[154] "The Impact of Model-Misspecification on Model Based Personalised Dosing".

[155] Holford, Ma, and Metz, "TDM Is Dead. Long Live TCI!"

ISoP recommends industry should demonstrate a positive benefit-risk in every individual patient, not just on summary statistics. For now, this is left largely to academic efforts (as shown in section 1.4.3), who need to start from scratch, as clinical trial results from drug development are generally not available, even if this is a legal requirement[156][157].

Tacrolimus still lacks a formally established exposure-response association for efficacy and safety, even though it was approved by FDA for use in liver transplant patients more than 25 years ago. Different centers use different PK target ranges, even discussing whether trough or AUC sampling is more appropriate. Some even consider whole-blood sampling inappropriate, advocating sampling at the transplanted organ instead. Furthermore, the effect is confounded by other immunosuppressive drugs such as prednisolone and mycophenolate mofetil (MMF), which are co-administered to ensure sufficient immunosuppression during potential tacrolimus underexposure, and are discontinued when stable tacrolimus exposure is reached some time after transplant. As MIPD is predicted to increase probability of target attainment, this may allow reduced prednisolone or mycophenolic acid doses, and even allow identification of the tacrolimus exposure-response relationship in patients.

In general, investigating exposure-response is complicated by multiple levels of inter-patient and within-patient variability on both PK and PD. This may be more efficiently investigated through the use of concentration-controlled trials (CCT), targeting specific exposures rather than dose levels[158]. A search on clinicaltrials.gov for the term "concentration controlled" yielded 37 results, of which 32 were in immunosuppressive treatment for transplant patients. It seems there is little incentive to perform a CCT during drug development if not strictly necessary, and regulators are wary to make market access even more difficult. Although we claim to use model-informed

[156] Overall, trial results are not reported in 1 out of every 4 studies. Only summary statistics are reported; raw data is rarely available and protected by both intellectual property and data privacy.

[157] Goldacre et al., "Compliance with Requirement to Report Results on the EU Clinical Trials Register".

[158] Sanathanan and Peck, "The Randomized Concentration-Controlled Trial".

drug development, trials with fixed dose groups are easier to execute, even if they yield less powerful results.

10.3. Simulation

Many authors have theorized and written viewpoints, commentaries and reviews[159] exclaiming their disappointment at the poor use of model-informed precision dosing, shouting ever louder for roadblocks to be lifted. Through simulation, we offer a much-needed perspective to counter this enthousiasm that is founded in theory but not reflected in clinical practice.

Roadblocks to precision dosing will not be lifted on fervour alone. We have identified roadblocks[160], yet strong arguments are required to justify investment into additional bio-analysis assays, better models for individual exposure-response, and userfriendly software adapted to fit the clinical workflow with minimal disruption. Through simulation, we have demonstrated that precision dosing of infliximab induction therapy does not justify this investment. In strong contrast, we demonstrated that tacrolimus MIPD not only improves target attainment, but also provides faster target attainment and reduces outliers. **We have demonstrated simulations can inform investment decisions, removing roadblocks where it matters.**

Wright et al rightfully louded the work of colleagues publishing the encouraging clinical trial results of PK-guided paclitaxel dosing. However, for every published positive study, 4 negative studies are left unpublished.[161] As demonstrated in section 1.4.3, the design of these trials leaves much to be desired. Statistical power is based on guesswork, further assuming normality in non-normal-distributed endpoints. For tacrolimus, we generated a

[159] Scheetz et al., "The Case for Precision Dosing"; Darwich et al., "Why Has Model-Informed Precision Dosing Not Yet Become Common Clinical Reality?"; Polasek, Shakib, and Rostami-Hodjegan, "Precision Dosing in Clinical Medicine"; Polasek et al., "Toward Dynamic Prescribing Information"; Maxfield and Zineh, "Precision Dosing"; Wright, Martin, and Cremers, "Spotlight Commentary"; Peck, "Precision Dosing".

[160] Darwich et al., "Why Has Model-Informed Precision Dosing Not Yet Become Common Clinical Reality?"

[161] Hopewell et al., "Publication Bias in Clinical Trials Due to Statistical Significance or Direction of Trial Results".

hypothetical in silico dataset for the MIPD arm, allowing targeted clinical trial design.

By providing a clear blueprint and supporting tool, and demonstrating this approach on several usecases, we have made simulation of MIPD more accessible. Where ardour and zeal failed to leverage the required additional investment in precision dosing, simulation can provide the rational arguments for MIPD implementation.

10.4. Implementation

Demonstrating clearly that tacrolimus MIPD was worth the investment, a software tool for precision dosing was implemented. By its open-source nature, this software tool managed to integrate almost seamlessly with current clinical workflow.

Three caveats became painfully apparent in this approach. To implement precision dosing in an existing clinical practice, both software tool and physician workflow have to be in sync. With commercial software, the physician is generally required to adapt. This hinders adoption. Optionally, the hospital may pay the software vendor to adapt or integrate the software. With open-source software, any commercial developer may adapt existing software to fit a specific clinical workflow. For precision dosing, the skillset required in these developers is very specific however, so this route may not always be feasible.

A second major roadblock is the legal climate surrounding medical devices. While an update of medical device legislation was urgently needed to provide a legal framework for mobile apps, smart watches, measuring scales, smoothie makers, exercise bikes, and many other gadgets, this legislation created a veritable maze for an open-source precision dosing framework. Ironically, the power of open-source development has long since been recognized in software development. More recently, the open-source model was used to develop ventilators or face masks to combat the COVID-19 pandemic[162]. Industry and regulators frown upon these initiatives however,

[162] Frazer, Shard, and Herdman, "Involvement of the Open-Source Community in Combating the Worldwide COVID-19 Pandemic"; Pearce, "A Review of Open Source Ventilators for COVID-19 and Future Pandemics".

either silently tolerating their use, or only approving use when no other alternatives are available[163].

Finally, there is the issue of cost. In most countries, precision dosing is not considered a reimbursable cost. Belgian nomenclature of health care actions does not provide any way to register this activity. Why would a physician or pharmacist provide this extra service, at considerable personal investment in both education and software tools, if the potential benefit for patients is not recognized by the health care payer? Moreover, any deviation from the label, even if founded in established science, is considered off-label use. The reimbursement of Remicade® (infliximab) for the treatment of ulcerative colitis is governed by K.B. 01.02.2018 - IV - 3960000 in Belgium, which prescribes an induction regimen of 5mg/kg at week 0, 2, and 6. Even when a growing body of scientific evidence justifies dose escalation, health care payers may not follow suit, potentially exploding costs for individual patients exhibiting unfortunate pharmacokinetics (e.g. high distribution volume or high drug clearance).

This situation may be solved at three points during market access. One way is to raise the bar for market access -as suggested by ISoP[164]-, requiring manufacturers to investigate precision dosing. Fixed dosing would only be accepted if precision dosing is demonstrated to not markedly improve outcomes. Alternatively, drugs could be approved at a candidate exposure, accompanied by a *suggested* posology to attain this exposure[165]. Doctors should be free to deviate from suggested dosing if evidence is available to support this. Finally, health care payers may negotiate value-based pricing[166],

---

163 https://www.gov.uk/government/publications/medical-devices-given-exceptional-use-authorisations-during-the-covid-19-pandemic, consulted 23-DEC-2021

164 Maloney et al., "Comment from International Society of Pharmacometrics on Exposure-Response Analysis in Drug Development and Regulatory Decision Making; Request for Comments (Docket No. FDA-2018-N-0791)".

165 Michael Neely, "Are We Really Going to Buy Into Individualized Dosing?", FDA workshop on Precision Dosing 2019, available at https://www.fda.gov/media/130406/download

166 Garrison and Towse, "Value-Based Pricing and Reimbursement in Personalised Healthcare".

forcing manufacturers and health care providers to collaborate in achieving optimal exposure and outcomes, possibly at individualized doses.

Regulators should consider the use of value-based pricing to incentivise optimal dosing and optimal exposure-efficacy-safety for every individual patient. Market access is currently based on a selected (generally, one-dose-fits-all) posology that was demonstrated in drug development to have acceptable efficacy/safety in the general population. The manufacturer then establishes a set price per dose with the health care payer. This is generally a country government, who sets guidelines for compensation of the drug to health care providers. Zooming in on infliximab (Remicade® by Johnson & Johnson) in Belgium (see also Chapter 1), a single vial costs €375 [167], of which €320 is paid by the health care payer RIZIV in Belgium. Currently, even if our model recommends a dose increase to 10mg/kg, this would cost patients an additional €1,200 for induction therapy.

Adapting and maintaining a software tool incurs a development cost as well. This should not be an insurmountable barrier, as we can leverage previous experience introducing software at general practitioners in Belgium[168]. Overall health benefits (and reduction in medicine reimbursement costs) should be relatively easy to demonstrate in a health economic study. Further collecting real-world evidence[169] in a continuous feedback loop[170] should serve to further improve predictive models and further refine dose adaptation.

[167] RIZIV CTGRM database, consulted on 23-DEC-21

[168] Van der Stighelen et al., "ICT-Tools En Gebruik van EMD Door de Huisarts".

[169] Radawski et al., "The Utility of Real-world Evidence for Benefit-risk Assessment, Communication, and Evaluation of Pharmaceuticals".

[170] Ribba et al., "Model-informed Artificial Intelligence".

10.5. Key points

Rational development of model-informed precision dosing is urgently needed. We demonstrated a simulation approach can lead to better models, evidence for precision dosing benefit, and better clinical trials. A software framework allows flexible implementation of precision dosing software adapted to physician workflow. This paves the road toward precision dosing, allowing rational investment supported by quantitative predictions.

Scientific acknowledgements

Overall: This work was funded by FWO grant TBM T003117N. We thank Alain Deleener for supervising the project at FWO. We thank Birgit Peeters for her support. We thank Dirk Kuypers, Geert Verbeke and Minne Casteels for their role in the management committee of this project.

This work was partly funded by the Department of Pharmaceutical and Pharmacological Sciences at the University of Leuven.

Ruben Faelens wrote the text. Pieter Annaert, Dirk Kuypers and Daniel Röshammar critically revised the text.

Chapter 1 Thomas Bouillon and Ruben Faelens designed the research. Thomas Bouillon, Daniel Röshammar, Dirk Kuypers and Pieter Annaert critically revised the text.

0 We thank Rachel Gregory, Ilse Sienaert and Ivo Roelants from LRD; Frank Hulstaert from KCE; Duquesne Julie and Alexandre Jauniaux from FAGG; Anton Vedder from the Centre for IT and Intellectual Property for their guidance regarding medical device legislation surrounding the open-sourcing of tdmore. We thank Egon Nijns, Bart Decuypere, Thomas Noppe and Bart Van den Bosch for development support at NexuzHealth/UZ Leuven. We thank Nicolas Luyckx and Quentin Leirens from SGS Exprimo for development support. We thank the Data Protection Officer for critically reviewing GDPR requirements. We thank all attendees on the tdmore workshops in 2019 and 2021.

Chapter 1 Conceptualization, Ruben Faelens, Thomas Bouillon and Erwin Dreesen; Data curation, Paul Declerck, Marc Ferrante, Séverine Vermeire and Erwin Dreesen; Formal analysis, Ruben Faelens and Zhigang Wang; Funding acquisition, Erwin Dreesen; Investigation, Ruben Faelens, Zhigang Wang and Erwin Dreesen; Methodology, Ruben Faelens, Zhigang Wang and Erwin Dreesen; Project administration, Erwin Dreesen; Resources, Paul Declerck, Marc Ferrante, Séverine Vermeire and Erwin Dreesen; Software, Ruben Faelens; Supervision, Thomas Bouillon and Erwin Dreesen; Validation, Ruben Faelens, Thomas Bouillon and Erwin Dreesen; Visualization, Ruben Faelens, Zhigang Wang and Erwin Dreesen; Writing – original draft, Ruben Faelens,

Zhigang Wang and Erwin Dreesen; Writing – review & editing, Thomas Bouillon, Paul Declerck, Marc Ferrante and Séverine Vermeire.

We acknowledge Vera Ballet for an excellent job in maintaining the Leuven IBD patient database; Griet Compernolle, Sophie Tops, Els Brouwers, and Miet Peeters for performing the infliximab serum concentration measurements; and Pieter Annaert for the guidance and management of PhD student Ruben Faelens. We thank Ann Gils for her guidance and supervision early in this project, and the UZLeuven IBD group for their critical review of early simulation results.

Chapter 7 Ruben Faelens, Pieter Annaert and Dirk Kuypers wrote the manuscript; Ruben Faelens, Dirk Kuypers and Thomas Bouillon designed the research; Ruben Faelens and Thomas Bouillon performed the research; Ruben Faelens analyzed the data; Ruben Faelens and Nicolas Luyckx contributed new analytical tools.

The authors wish to acknowledge the contributions of Dr. Thomas Van Hove in collecting the dataset, Ir. Quentin Leirens in supporting the development of the software, and Dr. Daniel Röshammar for reviewing the manuscript. We recognize the support of Annouschka Laenen at the KULeuven Biostats department in providing feedback to the study statistical analysis plan.

Chapter 1 We thank Annouschka Laenen for critical review of the statistical analysis plan, on which this work is based.

Chapter 1 Mirthe Vincken performed the research. Ruben Faelens designed the research, provided guidance, and wrote the manuscript.

Personal contribution

The author of this manuscript fully contributed to the design of the studies, the generation and assembly of data, the analysis and interpretation of the data, the design, implementation and testing of the software, the deployment and follow-up of clinical studies, and wrote all text contained in this manuscript, except for the following: Zhigang Wang wrote the Introduction of Chapter 1, and Erwin Dreesen partly wrote the Discussion of Chapter 1.

Conflict of interest

From 2017 to 2019, Ruben Faelens was an employee of SGS Exprimo. In this function, Ruben Faelens performed consulting work for Astellas Pharma in 2017 as a paid consultant on an unrelated compound in clinical development, and for Janssen Pharmaceutica on compounds not related to Remicade. From February 2022 onwards, Ruben was an employee of Janssen Pharmaceutica.

Thomas Bouillon was an employee of BioNotus GCV from 2019 to present.

Daniel Röshammar was an employee of SGS Exprimo (2016 to 2018) and Pharmetheus (2018 to present), performing paid consultancy for pharmaceutical companies on compounds unrelated to the work presented in this manuscript.

M.F. received research grant from AbbVie, Amgen, Biogen, Janssen, Pfizer, and Takeda; speakers fee from Abbvie, Amgen, Biogen, Boehringer-Ingelheim, Falk, Ferring, Janssen, Lamepro, MSD, Mylan, Pfizer, Sandoz, Takeda and Truvion Healthcare; consultancy fee from Abbvie, Boehringer-Ingelheim, Celltrion, Janssen, Lilly, Medtronic, MSD, Pfizer, Sandoz, Takeda and Thermo Fisher. S.V. received financial support for research from AbbVie, J&J, Pfizer, and Takeda; consulting and/or speaking fees from AbbVie, Arena Pharmaceuticals, Avaxia, Boehringer Ingelheim, Celgene, Dr Falk Pharma, Ferring, Galapagos, Genentech-Roche, Gilead, Hospira, Janssen, Mundipharma, MSD, Pfizer, Prodigest, Progenity, Prometheus, Robarts Clinical Trials, Second Genome, Shire, Takeda, Theravance, and Tillots Pharma AG. E.D. received consultancy fees from argenx and Janssen (all fees paid to the University). Z.W., and P.D. declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Curriculum Vitae and publication list

PERSONAL DATA

December 11, 1987
Belgian, Male
Wijgmaal, Belgium
+32 494 06 72 59
ruben.faelens@gmail.com
be.linkedin.com/in/rfaelsn

Trilingual engineer, broad experience in model-informed drug development and software engineering. Bringing new insights through model simulation.

WORK EXPERIENCE

Assoc. Director — *Feb 2022–now* — Janssen Pharmaceutica

As a scientist at the Clinical Pharmacology and Pharmacometrics department, I support model-informed drug development by simulating how compounds behave in previously unstudied populations, formulations or dosing regimen. This makes drug development more efficient, reduces the need for clinical trials, and ultimately brings more drugs to market faster and cheaper.

PhD researcher — *2018–2022* — KU Leuven

Development of a flexible framework for model-informed precision dosing. Simulation of tacrolimus precision dosing in renal transplant recipients: popPK model development, simulation, clinical trial optimization. Implementation of bedside precision dosing tool, integration with EHR system and cloud deployment. Teaching at undergraduate (physics) and doctoral (simulation of non-linear mixed effects systems) level.

IT Scientist *2014–2018* SGS Exprimo

Scientific PK/PD modeling & simulation for pharma industry. Performing simulations based on scientific literature and/or custom-developed models. Clinical trial simulation and power calculation, bio-equivalence studies, model-informed precision dosing. Responsible for management of high
performance computing cluster, software development, computer systems validation.

Business development *2013* Altran

Ramping up fixed price IT project turnover. Technical sales, proposal writing, coaching of non-IT sales people, workload and budget estimation, business development strategy. Technical interview for new hires.

Technical Lead (as Altran Consultant) *2011–2013* Exprimo - Hoffman La Roche

PK/PD Modeling & Simulation software development. Monte Carlo simulation. Analysis of user requirements, software architecture, development. Coaching of junior developers. +600MD project, team of 4 people.

Technical Lead (as Altran Consultant) *2011–2012* GSK

Implementation of a web application for making Matlab models available through the web. Analysis of business context, functional/technical requirements. Definition of product roadmap. Implementation and support.
Coaching / training of other developers.

EDUCATION

PhD in pharmaceutical sciences (finishing) *2017 – now* Katholieke Universiteit Leuven

"Computer-assisted dosing framework and proof of concept in tacrolimus for renal transplant recipients."
Supervised by: P. ANNAERT, D. KUYPERS, D. ROSHAMMAR, T. BOUILLON. Funded by FWO-TBM grant.

Master of Engineering: Computer Science *2008–2010* Katholieke Universiteit Leuven

Master in Engineering · Computer Science: Distributed Systems
Thesis: “Using a java RDF model to improve the IDE environment” Advisor: Prof. Eric STEEGMANS (KUL) and Prof. Ambjorn NAEVE (KTH)

Erasmus exchange *Fall 2009* Kungliga Tekniska Högskolan, Stockholm

Worked on master thesis and some basic courses. Learned basic Swedish.

Bachelor of Engineering *2006-2008* Katholieke Universiteit Leuven

Bachelor of Engineering
Major: Computer Science · Minor: Electrical Engineering
Graduated with distinction

PUBLICATIONS AND POSTERS

1. Ruben Faelens, Nicolas Luyckx, Dirk Kuypers, Thomas Bouillon, and Pieter Annaert. Predicting model-informed precision dosing: a test-case in tacrolimus dose adaptation for kidney transplant recipients. *CPT: Pharmacometrics & Systems Pharmacology*, 2021. https://doi.org/10.1002/psp4.12758

2. Anne-Gaëlle Dosne, Elodie Valade, Kim Stuyckens, Peter De Porre, Anjali Avadhani, Anne O'Hagan, Lilian Y. Li, Daniele Ouellet, Ruben Faelens, Quentin Leirens, Italo Poggesi, and Juan Jose Perez Ruixo. Erdafitinib's effect on serum phosphate justifies its pharmacodynamically guided dosing in patients with cancer. *CPT: Pharmacometrics & Systems Pharmacology*, page psp4.12727, November 2021

3. Ruben Faelens, Nicolas Luyckx, Quentin Leirens, Thomas Bouillon, Dirk Kuypers, and Pieter Annaert. Building model-informed precision dosing software using R: blueprint for a state-of-the-art development process. In *PAGE 29 Conference poster*, 2021

4. Ruben Faelens, Zhigang Wang, Thomas Bouillon, Paul Declerck, Marc Ferrante, Séverine Vermeire, and Erwin Dreesen. Model-Informed Precision Dosing during Infliximab Induction Therapy Reduces Variability in Exposure and Endoscopic Improvement between Patients with Ulcerative Colitis. *Pharmaceutics*, 13(10):1623, October 2021

5. Ruben Faelens, Nicolas Luyckx, Quentin Leirens, Dirk Kuypers, and Thomas Bouillon. Model predictive control with Bayesian updates (MPC) is more robust to model misspecification, compared to standard Bayesian control (sEBE) for Therapeutic Drug Management (TDM). Investigation in a cohort of 315 patients receiving tacrolimus during the first 14d after renal transplantation. In *PAGE 28 Conference poster*, 2019

6. Anne-Gaelle Dosne, Elodie Valade, Kim Stuyckens, Peter De Porre, Loretta Sullivan-Chang,
Dominique Swerts, Daniele Ouellet, Lilian Li, Ruben Faelens, Quentin Leirens, and Juan Jose Perez Ruixo. Optimizing Biomarker-Based Dosing Algorithm for Erdafitinib through PK-PD Modeling and Simulations. In *PAGE 27 Conference poster*, 2018

7. Erwin Dreesen, Ruben Faelens, Gert A. Van Assche, Marc Ferrante, Severine Vermeire, Ann Gils, and Thomas Bouillon. Mo1846 - Adaptive Dosing During Infliximab Induction Therapy can Improve Mucosal Healing Rates in Patients with Ulcerative Colitis. *Gastroenterology*, 154(6):S–823, May 2018

8. E Dreesen, R Faelens, G Van Assche, M Ferrante, S Vermeire, A Gils, and T Bouillon. P642 Adaptive dosing during infliximab induction therapy can improve mucosal healing rates in patients with ulcerative colitis. *Journal of Crohn's and Colitis*, 12(supplement_1):S434–S435, January 2018

9. Ruben Faelens, Erwin Dreesen, Gert Van Assche, Marc Ferrante, Séverine Vermeire, Ann Gils, and Thomas Bouillon. Benefits of TDM in clinical management: a test case in infliximab for ulcerative colitis. In *PAGE 27 Conference poster*, 2018

10. C. Tang, N. Knops, R. Faelens, D. Kuypers, and T. Bouillon. Population Pharmacokinetics of Tacrolimus in Pediatric Patients having undergone kidney, liver or bowel/liver transplantation. In *PAGE 27 Conference poster*, 2018

11. Erwin Dreesen, Ruben Faelens, Gert Van Assche, Marc Ferrante, Séverine Vermeire, Ann Gils, and Thomas Bouillon. Development of a population pharmacokinetic and pharmacodynamic model to describe the effect of infliximab induction therapy on mucosal healing in patients with ulcerative colitis. In *PAGE 27 Conference poster*, 2018

12. Quentin Leirens, Belen Valenzuela, Andreas Lindauer, Ruben Faelens, Per Olsson, and Daniel Roeshammar. Clinical Trial Simulation of a Phase I Paediatric Oncology Study using Simulo. In *PAGE 27 Conference poster*, 2018

13. Quentin Leirens, Ruben Faelens, Per Olsson Gisleskog, Andreas Lindauer, and Daniel Roshammar. Simulo – a platform for advanced model based simulations using R. In *PAGE 26 Software demonstration*, 2017

14. Ruben Faelens, Philippe Jacqmin, Per Olsson Gisleskog, Andreas Lindauer, and Daniel Roeshammar. Clinical trial optimization of efficacy studies in slowly progressive diseases. In *PAGE 26 Conference poster*, 2017

15. J. Llaudó, L. Anta, I. Ayani, J. Martínez-González, I. Gutierro, R. Faelens, J. Winkler, and E. Snoeck. Population pharmacokinetic modeling and simulations of dopamine Dd2 receptor occupancy of long-acting intramuscular risperidone-ISM. *European Psychiatry*, 33(S1):S572–S572, March 2016

16. Ruben Faelens, Quentin Leirens, and Philippe Jacqmin. Simulo: a new PK-PD-Disease model simulator. In *PAGE 24 Software demonstration*, 2015

References

Abrantes, João A., Siv Jönsson, Mats O. Karlsson, and Elisabet I. Nielsen. "Handling Inter-occasion Variability in Model-based Therapeutic Drug Monitoring." In *Annual Meeting of the Population Approach Group in Europe, Budapest, Hungary*, Vol. 6, 2017.

Adedokun, Omoniyi J., William J. Sandborn, Brian G. Feagan, Paul Rutgeerts, Zhenhua Xu, Colleen W. Marano, Jewel Johanns, et al. "Association Between Serum Concentration of Infliximab and Efficacy in Adult Patients With Ulcerative Colitis." *Gastroenterology* 147, no. 6 (December 2014): 1296–1307.e5. https://doi.org/10.1053/j.gastro.2014.08.035.

Ahluwalia, Arti, Carmelo De Maria, Andrés Díaz Lantada, Licia Di Pietro, Alice Ravizza, Mannan Mridha, June Madete, et al. "Towards Open Source Medical Devices:" In *Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies*, 141–49. Funchal, Madeira, Portugal: SCITEPRESS - Science; Technology Publications, 2018. https://doi.org/10.5220/0006586501410149.

Alihodzic, Dzenefa, Astrid Broeker, Michael Baehr, Stefan Kluge, Claudia Langebrake, and Sebastian Georg Wicha. "Impact of Inaccurate Documentation of Sampling and Infusion Time in Model-Informed Precision Dosing." *Frontiers in Pharmacology* 11 (March 2020): 172. https://doi.org/10.3389/fphar.2020.00172.

Anderson, J. E., A. S. Munday, A. W. Kelman, B. Whiting, J. D. Briggs, J. Knepil, and A. H. Thomson. "Evaluation of a Bayesian Approach to the Pharmacokinetic Interpretation of Cyclosporin Concentrations in Renal Allograft Recipients:" *Therapeutic Drug Monitoring* 16, no. 2 (April 1994): 160–65. https://doi.org/10.1097/00007691-199404000-00009.

Andreu-Perez, Javier, Daniel R. Leff, H. M. D. Ip, and Guang-Zhong Yang. "From Wearable Sensors to Smart Implants--Toward Pervasive and Personalized Healthcare." *IEEE Transactions on Biomedical Engineering* 62, no. 12 (2015): 2750–62. https://doi.org/10.1109/TBME.2015.2422751.

Antignac, Marie, Benoit Barrou, Robert Farinotti, Philippe Lechat, and Saïk Urien. "Population Pharmacokinetics and Bioavailability of Tacrolimus in Kidney Transplant Patients." *British Journal of Clinical Pharmacology* 64, no. 6 (April 2007): 750–57. https://doi.org/10.1111/j.1365-2125.2007.02895.x.

Arias, Maria Theresa, Niels Vande Casteele, Séverine Vermeire, Anthony de Buck van Overstraeten, Thomas Billiet, Filip Baert, Albert Wolthuis, et al. "A Panel to Predict Long-Term Outcome of Infliximab Therapy for Patients With Ulcerative Colitis."

Clinical Gastroenterology and Hepatology 13, no. 3 (March 2015): 531–38. https://doi.org/10.1016/j.cgh.2014.07.055.

Austin Health, and University of Melbourne. "Optimising Infliximab Induction Therapy for Acute Severe Ulcerative Colitis (PREDICT-UC)." In *Clinicaltrials.gov*. Nct02770040. Accessed July 14, 2021. https://clinicaltrials.gov/ct2/show/NCT02770040.

Balding, Philip R., Cristina S. Porro, Kirsty J. McLean, Michael J. Sutcliffe, Jean-Didier Maréchal, Andrew W. Munro, and Sam P. de Visser. "How Do Azoles Inhibit Cytochrome P450 Enzymes? A Density Functional Study." *The Journal of Physical Chemistry A* 112, no. 50 (December 2008): 12911–18. https://doi.org/10.1021/jp802087w.

Balzola, Federico, Charles Bernstein, Gwo Tzer Ho, and Charlie Lees. "Trough serum infliximab: A predictive factor of clinical outcome for infliximab treatment in acute ulcerative colitis: Commentary." *Inflammatory Bowel Disease Monitor* 59 (2009): 49–54.

Battat, Robert, Amy Hemperly, Stephanie Truong, Natalie Whitmire, Brigid S. Boland, Parambir S. Dulai, Ariela K. Holmer, et al. "Baseline Clearance of Infliximab Is Associated With Requirement for Colectomy in Patients With Acute Severe Ulcerative Colitis." *Clinical Gastroenterology and Hepatology* 19, no. 3 (March 2021): 511–518.e6. https://doi.org/10.1016/j.cgh.2020.03.072.

Bauer, Robert J. "NONMEM Tutorial Part I: Description of Commands and Options, with Simple Examples of Population Analysis." *CPT: Pharmacometrics & Systems Pharmacology* 8, no. 8 (2019): 525–37.

———. "NONMEM Tutorial Part II: Estimation Methods and Advanced Examples." *CPT: Pharmacometrics & Systems Pharmacology* 8, no. 8 (2019): 538–56.

Bégaud, B. "The Rennes Disaster: Lessons to Be Drawn." *Clinical Therapeutics* 39, no. 8 (2017): e110.

Bejan-Angoulvant, Theodora, David Ternant, Fadela Daoued, Frédéric Medina, Louis Bernard, Saloua Mammou, Gilles Paintaud, and Denis Mulleman. "Brief Report: Relationship Between Serum Infliximab Concentrations and Risk of Infections in Patients Treated for Spondyloarthritis." *Arthritis & Rheumatology* 69, no. 1 (January 2017): 108–13. https://doi.org/10.1002/art.39841.

Ben-Horin, Shomron, Uri Kopylov, and Yehuda Chowers. "Optimizing Anti-TNF Treatments in Inflammatory Bowel Disease." *Autoimmunity Reviews* 13, no. 1 (2014): 24–30. https://doi.org/https://doi.org/10.1016/j.autrev.2013.06.002.

Bergmann, Troels K., Stefanie Hennig, Katherine A. Barraclough, Nicole M. Isbel, and Christine E. Staatz. "Population Pharmacokinetics of Tacrolimus in Adult Kidney

Transplant Patients: Impact of Cyp3a5 Genotype on Starting Dose." *Therapeutic Drug Monitoring* 36, no. 1 (February 2014): 62–70. https://doi.org/10.1097/FTD.0b013e31829f1ab8.

Bouamar, R., N. Shuker, D. A. Hesselink, W. Weimar, H. Ekberg, B. Kaplan, C. Bernasconi, and T. van Gelder. "Tacrolimus Predose Concentrations Do Not Predict the Risk of Acute Rejection After Renal Transplantation: A Pooled Analysis from Three Randomized-Controlled Clinical Trials(†)." *American Journal of Transplantation: Official Journal of the American Society of Transplantation and the American Society of Transplant Surgeons* 13, no. 5 (May 2013): 1253–61. https://doi.org/10.1111/ajt.12191.

Brandse, Johannan F., Ron A. Mathôt, Desiree van der Kleij, Theo Rispens, Yaël Ashruf, Jeroen M. Jansen, Svend Rietdijk, et al. "Pharmacokinetic Features and Presence of Antidrug Antibodies Associate With Response to Infliximab Induction Therapy in Patients With Moderate to Severe Ulcerative Colitis." *Clinical Gastroenterology and Hepatology* 14 (2016): 251, e1-2-258. https://doi.org/10.1016/j.cgh.2015.10.029.

Brekkan, Ari, Luis Lopez-Lazaro, Gunnar Yngman, Elodie L. Plan, Chayan Acharya, Andrew C. Hooker, Suresh Kankanwadi, and Mats O. Karlsson. "A Population Pharmacokinetic-Pharmacodynamic Model of Pegfilgrastim." *The AAPS Journal* 20, no. 5 (September 2018): 91. https://doi.org/10.1208/s12248-018-0249-y.

Broeker, A. "Towards Precision Dosing of Vancomycin: A Systematic Evaluation of Pharmacometric Models for Bayesian Forecasting." *Clinical Microbiology and Infection*, 2019, 7.

Brooks, Emily, Susan E. Tett, Nicole M. Isbel, and Christine E. Staatz. "Population Pharmacokinetic Modelling and Bayesian Estimation of Tacrolimus Exposure: Is This Clinically Useful for Dosage Prediction Yet?" *Clinical Pharmacokinetics* 55, no. 11 (November 2016): 1295–335. https://doi.org/10.1007/s40262-016-0396-1.

Brunet, Mercè, Teun van Gelder, Anders Åsberg, Vincent Haufroid, Dennis A Hesselink, Loralie Langman, Florian Lemaitre, et al. "Therapeutic Drug Monitoring of Tacrolimus-Personalized Therapy: Second Consensus Report." *Ther Drug Monit* 41, no. 3 (2019): 47.

Buclin, Thierry, Yann Thoma, Nicolas Widmer, Pascal André, Monia Guidi, Chantal Csajka, and Laurent A. Decosterd. "The Steps to Therapeutic Drug Monitoring: A Structured Approach Illustrated with Imatinib." *Frontiers in Pharmacology* 11 (2020): 177.

Bultheel, Adhemar. *Inleiding Tot de Numerieke Wiskunde*. Leuven: Acco, 2007.

Campagne, Olivia, Donald E. Mager, and Kathleen M. Tornatore. "Population Pharmacokinetics of Tacrolimus in Transplant Recipients: What Did We Learn About

Sources of Interindividual Variabilities?" *The Journal of Clinical Pharmacology* 59, no. 3 (March 2019): 309–25. https://doi.org/10.1002/jcph.1325.

Chheda, Jason J., Andrew Tarleton, and Jay H. Eidem. "Targeted Aspergillus Prophylaxis With Voriconazole in Heart Transplant Patients: A Focus on the Interaction With Tacrolimus." *Journal of Pharmacy Technology* 35, no. 4 (2019): 164–71.

D'Haens, Geert, Severine Vermeire, Guy Lambrecht, Filip Baert, Peter Bossuyt, Benjamin Pariente, Anthony Buisson, et al. "Increasing Infliximab Dose Based on Symptoms, Biomarkers, and Serum Drug Concentrations Does Not Increase Clinical, Endoscopic, and Corticosteroid-Free Remission in Patients With Active Luminal Crohn's Disease." *Gastroenterology* 154, no. 5 (April 2018): 1343–1351.e1. https://doi.org/10.1053/j.gastro.2018.01.004.

Darwich, A S, K Ogungbenro, A A Vinks, J R Powell, J-L Reny, N Marsousi, Y Daali, et al. "Why Has Model-Informed Precision Dosing Not Yet Become Common Clinical Reality? Lessons From the Past and a Roadmap for the Future." *Clinical Pharmacology & Therapeutics* 101, no. 5 (May 2017): 646–56. https://doi.org/10.1002/cpt.659.

Dickson, Rachael, Jessica Bell, Amber Dar, Laura Downey, Muireann Quigley, and Victoria Moore. "#WeAreNotWaiting DIY Artificial Pancreas Systems and Challenges for the Law." *Diabetic Medicine*, October 2021. https://doi.org/10.1111/dme.14715.

Ding, N. S., A. Hart, and P. De Cruz. "Systematic Review: Predicting and Optimising Response to Anti-TNF Therapy in Crohn's Disease - Algorithm for Practical Management." *Alimentary Pharmacology & Therapeutics* 43, no. 1 (January 2016): 30–51. https://doi.org/10.1111/apt.13445.

Dosne, Anne-Gaëlle, Elodie Valade, Kim Stuyckens, Peter De Porre, Anjali Avadhani, Anne O'Hagan, Lilian Y. Li, et al. "Erdafitinib's Effect on Serum Phosphate Justifies Its Pharmacodynamically Guided Dosing in Patients with Cancer." *CPT: Pharmacometrics & Systems Pharmacology*, November 2021, psp4.12727. https://doi.org/10.1002/psp4.12727.

Dreesen, E., R. Faelens, G. Van Assche, M. Ferrante, S. Vermeire, A. Gils, and T. Bouillon. "Optimising Infliximab Induction Dosing for Patients with Ulcerative Colitis." *British Journal of Clinical Pharmacology* 85, no. 4 (2019). https://doi.org/10.1111/bcp.13859.

Dreesen, Erwin. "New Tools for Therapeutic Drug Monitoring: Making Big Things Out of Small Pieces." *Journal of Crohn's and Colitis*, August 2021, jjab137. https://doi.org/10.1093/ecco-jcc/jjab137.

Dreesen, Erwin, Sophie Berends, David Laharie, Geert D'Haens, Séverine Vermeire, Ann Gils, and Ron Mathôt. "Modelling of the Relationship Between Infliximab

Exposure, Faecal Calprotectin and Endoscopic Remission in Patients with Crohn's Disease." *British Journal of Clinical Pharmacology* 87, no. 1 (January 2021): 106–18. https://doi.org/10.1111/bcp.14364.

Dreesen, Erwin, Ruben Faelens, Gert Van Assche, Marc Ferrante, Séverine Vermeire, Ann Gils, and Thomas Bouillon. "Optimising infliximab induction dosing for patients with ulcerative colitis." *British Journal of Clinical Pharmacology* 85, no. 4 (2019): 782–95. https://doi.org/10.1111/bcp.13859.

Faelens, Ruben, Nicolas Luyckx, and Quentin Leirens. "Tdmore." KULeuven, September 2020. github.com/tdmore-dev/tdmore.

Farkas, Klaudia, Mariann Rutka, Petra A. Golovics, Zsuzsanna Végh, Barbara D. Lovász, Tibor Nyári, Krisztina B. Gecse, et al. "Efficacy of infliximab biosimilar CT-P13 induction therapy on mucosal healing in ulcerative colitis." *Journal of Crohn's and Colitis* 10, no. 11 (2016): 1273–78. https://doi.org/10.1093/ecco-jcc/jjw085.

Fidler, Matthew L., Melissa Hallow, Justin Wilkins, and Wenping Wang. *RxODE: Facilities for Simulating from ODE-Based Models*, 2021. https://CRAN.R-project.org/package=RxODE.

Fidler, Matthew, Justin J. Wilkins, Richard Hooijmaijers, Teun M. Post, Rik Schoemaker, Mirjam N. Trame, Yuan Xiong, and Wenping Wang. "Nonlinear Mixed-Effects Model Development and Simulation Using Nlmixr and Related R Open-Source Packages." *CPT: Pharmacometrics & Systems Pharmacology* 8, no. 9 (September 2019): 621–33. https://doi.org/10.1002/psp4.12445.

Frazer, John Scott, Amelia Shard, and James Herdman. "Involvement of the Open-Source Community in Combating the Worldwide COVID-19 Pandemic: A Review." *Journal of Medical Engineering & Technology* 44, no. 4 (May 2020): 169–76. https://doi.org/10.1080/03091902.2020.1757772.

Garrison, Louis, and Adrian Towse. "Value-Based Pricing and Reimbursement in Personalised Healthcare: Introduction to the Basic Health Economics." *Journal of Personalized Medicine* 7, no. 3 (September 2017): 10. https://doi.org/10.3390/jpm7030010.

Gelman, Andrew. *Bayesian Data Analysis*. Third edition. Chapman & Hall/CRC Texts in Statistical Science. Boca Raton: CRC Press, 2014.

Goldacre, Ben, Nicholas J DeVito, Carl Heneghan, Francis Irving, Seb Bacon, Jessica Fleminger, and Helen Curtis. "Compliance with Requirement to Report Results on the EU Clinical Trials Register: Cohort Study and Web Resource." *BMJ*, September 2018, k3218. https://doi.org/10.1136/bmj.k3218.

Golubović, Bojana, Katarina Vučićević, Dragana Radivojević, Sandra Vezmar Kovačević, Milica Prostran, and Branislava Miljković. "Total Plasma Protein Effect on

Tacrolimus Elimination in Kidney Transplant Patients – Population Pharmacokinetic Approach." *European Journal of Pharmaceutical Sciences* 52 (February 2014): 34–40. https://doi.org/10.1016/j.ejps.2013.10.008.

Grosse Kathoefer, David, and Jens Leker. "Knowledge Transfer in Academia: An Exploratory Study on the Not-Invented-Here Syndrome." *The Journal of Technology Transfer* 37, no. 5 (October 2012): 658–75. https://doi.org/10.1007/s10961-010-9204-5.

Han, Nayoung, Soojung Ha, Hwi-yeol Yun, Myeong Gyu Kim, Sang-Il Min, Jongwon Ha, Jangik Ike Lee, Jung Mi Oh, and In-Wha Kim. "Population Pharmacokinetic-Pharmacogenetic Model of Tacrolimus in the Early Period After Kidney Transplantation." *Basic & Clinical Pharmacology & Toxicology* 114, no. 5 (May 2014): 400–406. https://doi.org/10.1111/bcpt.12176.

Han, Nayoung, Hwi-yeol Yun, Jin-yi Hong, In-Wha Kim, Eunhee Ji, Su Hyun Hong, Yon Su Kim, Jongwon Ha, Wan Gyoon Shin, and Jung Mi Oh. "Prediction of the Tacrolimus Population Pharmacokinetic Parameters According to Cyp3a5 Genotype and Clinical Factors Using NONMEM in Adult Kidney Transplant Recipients." *European Journal of Clinical Pharmacology* 69, no. 1 (January 2013): 53–63. https://doi.org/10.1007/s00228-012-1296-4.

Hastings, W. K. "Monte Carlo Sampling Methods Using Markov Chains and Their Applications." *Biometrika* 57, no. 1 (April 1970): 97–109. https://doi.org/10.1093/biomet/57.1.97.

Hedaya, Mohsen A. *Basic Pharmacokinetics*. CRC Press, 2012.

Hindmarsh, A. C., and L. R. Petzold. "LSODA, Ordinary Differential Equation Solver for Stiff or Non-Stiff System," 2005.

Holford, Nicholas H. G., and Thierry Buclin. "Safe and Effective Variability—A Criterion for Dose Individualization:" *Therapeutic Drug Monitoring* 34, no. 5 (October 2012): 565–68. https://doi.org/10.1097/FTD.0b013e31826aabc3.

Holford, Nick H. G., and Brian J. Anderson. "Allometric Size: The Scientific Theory and Extension to Normal Fat Mass." *European Journal of Pharmaceutical Sciences* 109 (November 2017): S59–64. https://doi.org/10.1016/j.ejps.2017.05.056.

Holford, Nick, Guangda Ma, and David Metz. "TDM Is Dead. Long Live TCI!" *British Journal of Clinical Pharmacology*, 2020.

Hopewell, Sally, Kirsty Loudon, Mike J Clarke, Andrew D Oxman, and Kay Dickersin. "Publication Bias in Clinical Trials Due to Statistical Significance or Direction of Trial Results." Edited by Cochrane Methodology Review Group. *Cochrane Database of Systematic Reviews* 2010, no. 1 (January 2009). https://doi.org/10.1002/14651858.MR000006.pub3.

Hughes, David M, Srijib Goswami, Ron J Keizer, Maria-Stephanie A Hughes, and Jonathan D Faldasz. “Bayesian Clinical Decision Support-Guided Versus Clinician-Guided Vancomycin Dosing in Attainment of Targeted Pharmacokinetic Parameters in a Paediatric Population.” *Journal of Antimicrobial Chemotherapy*, October 2019, dkz444. https://doi.org/10.1093/jac/dkz444.

Hughes, Jasmine H., and Ron J. Keizer. “A Hybrid Machine Learning/Pharmacokinetic Approach Outperforms Maximum a Posteriori Bayesian Estimation by Selectively Flattening Model Priors.” *CPT: Pharmacometrics & Systems Pharmacology*, July 2021, psp4.12684. https://doi.org/10.1002/psp4.12684.

Jusko, William J., Wojciech Piekoszewski, Goran B. Klintmalm, Mark S. Shaefer, Mary F. Hebert, Antoni A. Piergies, Charles C. Lee, Paul Schechter, and Qais A. Mekki. “Pharmacokinetics of Tacrolimus in Liver Transplant Patients*.” *Clinical Pharmacology & Therapeutics* 57, no. 3 (March 1995): 281–90. https://doi.org/10.1016/0009-9236(95)90153-1.

Kabir, Vincent, Johan Maertens, and Dirk Kuypers. “Fungal Infections in Solid Organ Transplantation: An Update on Diagnosis and Treatment.” *Transplantation Reviews* 33, no. 2 (April 2019): 77–86. https://doi.org/10.1016/j.trre.2018.12.001.

Kantasiripitak, Wannee, Ruth Van Daele, Matthias Gijsen, Marc Ferrante, Isabel Spriet, and Erwin Dreesen. “Software Tools for Model-Informed Precision Dosing: How Well Do They Satisfy the Needs?” *Frontiers in Pharmacology* 11 (May 2020): 620. https://doi.org/10.3389/fphar.2020.00620.

Kantasiripitak, Wannee, Bram Verstockt, Dahham Alsoud, Triana Lobatón, Debby Thomas, Ann Gils, Séverine Vermeire, Marc Ferrante, and Erwin Dreesen. “The Effect of Aging on Infliximab Exposure and Response in Patients with Inflammatory Bowel Diseases.” *British Journal of Clinical Pharmacology*, March 2021, bcp.14785. https://doi.org/10.1111/bcp.14785.

Karlsson, Mats O., Stuart L. Beal, and Lewis B. Sheiner. “Three New Residual Error Models for Population PK/PD Analyses.” *Journal of Pharmacokinetics and Biopharmaceutics* 23, no. 6 (1995): 651–72.

Keeler, Allison M., and Terence R. Flotte. “Recombinant Adeno-Associated Virus Gene Therapy in Light of Luxturna (and Zolgensma and Glybera): Where Are We, and How Did We Get Here?” *Annual Review of Virology* 6 (2019): 601–21.

Keizer, Ron J., Rob ter Heine, Adam Frymoyer, Lawrence J. Lesko, Ranvir Mangat, and Srijib Goswami. “Model-Informed Precision Dosing at the Bedside: Scientific Challenges and Opportunities.” *CPT: Pharmacometrics & Systems Pharmacology* 7, no. 12 (December 2018): 785–87. https://doi.org/10.1002/psp4.12353.

Kesavadev, Jothydev, Seshadhri Srinivasan, Banshi Saboo, Meera Krishna B, and Gopika Krishnan. “The Do-It-Yourself Artificial Pancreas: A Comprehensive Review.”

Diabetes Therapy 11, no. 6 (June 2020): 1217–35. https://doi.org/10.1007/s13300-020-00823-z.

Kobayashi, Taku, Yasuo Suzuki, Satoshi Motoya, Fumihito Hirai, Haruhiko Ogata, Hiroaki Ito, Noriko Sato, Kunihiko Ozaki, Mamoru Watanabe, and Toshifumi Hibi. "First trough level of infliximab at week 2 predicts future outcomes of induction therapy in ulcerative colitis—results from a multicenter prospective randomized controlled trial and its post hoc analysis." *Journal of Gastroenterology* 51 (2016): 241–51. https://doi.org/10.1007/s00535-015-1102-z.

Krieger, Alexandra, and Efstratios N. Pistikopoulos. "Model Predictive Control of Anesthesia Under Uncertainty." *Computers & Chemical Engineering* 71 (2014): 699–707. https://doi.org/https://doi.org/10.1016/j.compchemeng.2014.07.025.

Kuhn, E., and M. Lavielle. "Maximum Likelihood Estimation in Nonlinear Mixed Effects Models." *Computational Statistics & Data Analysis* 49, no. 4 (June 2005): 1020–38. https://doi.org/10.1016/j.csda.2004.07.002.

Kuypers, Dirk R. J., Kathleen Claes, Pieter Evenepoel, Bart Maes, and Yves Vanrenterghem. "Clinical Efficacy and Toxicity Profile of Tacrolimus and Mycophenolic Acid in Relation to Combined Long-Term Pharmacokinetics in de Novo Renal Allograft Recipients." *Clinical Pharmacology and Therapeutics* 75, no. 5 (May 2004): 434–47. https://doi.org/10.1016/j.clpt.2003.12.009.

Landemaine, Amandine, Antoine Petitcollin, Charlène Brochard, Céline Miard, Marie Dewitte, Eric Le Balc'h, Thomas Grainville, Eric Bellissant, Laurent Siproudhis, and Guillaume Bouguen. "Cumulative Exposure to Infliximab, But Not Trough Concentrations, Correlates With Rate of Infection." *Clinical Gastroenterology and Hepatology* 19, no. 2 (February 2021): 288–295.e4. https://doi.org/10.1016/j.cgh.2020.03.018.

Leander, Jacob, Joachim Almquist, Christine Ahlström, Johan Gabrielsson, and Mats Jirstrand. "Mixed Effects Modeling Using Stochastic Differential Equations: Illustrated by Pharmacokinetic Data of Nicotinic Acid in Obese Zucker Rats." *The AAPS Journal* 17, no. 3 (2015): 586–96.

Lesko, Lawrence J. "Perspective on Model-informed Drug Development." *CPT: Pharmacometrics & Systems Pharmacology*, August 2021, psp4.12699. https://doi.org/10.1002/psp4.12699.

Lindbom, Lars, Pontus Pihlgren, and Niclas Jonsson. "PsN-Toolkit—A Collection of Computer Intensive Statistical Methods for Non-Linear Mixed Effect Modeling Using NONMEM." *Computer Methods and Programs in Biomedicine* 79, no. 3 (September 2005): 241–57. https://doi.org/10.1016/j.cmpb.2005.04.005.

Magro, Fernando, Paolo Gionchetti, Rami Eliakim, Sandro Ardizzone, Alessandro Armuzzi, Manuel Barreiro-de Acosta, Johan Burisch, et al. "Third European Evidence-

Based Consensus on Diagnosis and Management of Ulcerative Colitis. Part 1: Definitions, Diagnosis, Extra-Intestinal Manifestations, Pregnancy, Cancer Surveillance, Surgery, and Ileo-Anal Pouch Disorders." *Journal of Crohn's and Colitis* 11, no. 6 (June 2017): 649–70. https://doi.org/10.1093/ecco-jcc/jjx008.

Maloney, Al, Nick Holford, Joachim Grevel, David Norris, Amit Taneja, Mariam Ahmed, Thierry Buclin, and Jonathan French. "Comment from International Society of Pharmacometrics on Exposure-Response Analysis in Drug Development and Regulatory Decision Making; Request for Comments (Docket No. FDA-2018-N-0791)," July 2018. https://www.regulations.gov/comment/FDA-2018-N-0791-0020.

Martin, Andrew D., Kevin M. Quinn, and Jong Hee Park. "**MCMCpack** : Markov Chain Monte Carlo in *r*." *Journal of Statistical Software* 42, no. 9 (2011). https://doi.org/10.18637/jss.v042.i09.

Maxfield, Kimberly, and Issam Zineh. "Precision Dosing: A Clinical and Public Health Imperative." *JAMA* 325, no. 15 (April 2021): 1505. https://doi.org/10.1001/jama.2021.1004.

McDougall, David AJ, Jennifer Martin, E. Geoffrey Playford, and Bruce Green. "The Impact of Model-Misspecification on Model Based Personalised Dosing." *The AAPS Journal* 18, no. 5 (2016): 1244–53.

Miano, Todd A., Judd D. Flesch, Rui Feng, Caitlin M. Forker, Melanie Brown, Michelle Oyster, Laurel Kalman, et al. "Early Tacrolimus Concentrations After Lung Transplant Are Predicted by Combined Clinical and Genetic Factors and Associated With Acute Kidney Injury." *Clinical Pharmacology & Therapeutics* 107, no. 2 (February 2020): 462–70. https://doi.org/10.1002/cpt.1629.

Mitrev, N., N. Vande Casteele, C. H. Seow, J. M. Andrews, S. J. Connor, G. T. Moore, M. Barclay, et al. "Review article: consensus statements on therapeutic drug monitoring of anti-tumour necrosis factor therapy in inflammatory bowel diseases." *Alimentary Pharmacology and Therapeutics* 46 (2017): 1037–53. https://doi.org/10.1111/apt.14368.

Monchaud, Caroline, Brenda C. De Winter, Christiane Knoop, Marc Estenne, Martine Reynaud-Gaubert, Christophe Pison, Marc Stern, et al. "Population Pharmacokinetic Modelling and Design of a Bayesian Estimator for Therapeutic Drug Monitoring of Tacrolimus in Lung Transplantation." *Clinical Pharmacokinetics*, 2012. https://doi.org/10.2165/11594760-000000000-00000.

"Monolix Version 2019r1." Lixoft SAS, n.d. http://lixoft.com/products/monolix/.

Musuamba, Flora T., Michel Mourad, Vincent Haufroid, Martine De Meyer, Arnaud Capron, Isabelle K. Delattre, Roger K. Verbeeck, and Pierre Wallemacq. "Statistical Tools for Dose Individualization of Mycophenolic Acid and Tacrolimus Co-Administered During the First Month After Renal Transplantation." *British Journal of*

Clinical Pharmacology 75, no. 5 (2013): 1277–88. https://doi.org/https://doi.org/10.1111/bcp.12007.

Musuamba, Flora Tshinanu, Michel Mourad, Vincent Haufroid, Martine Demeyer, Arnaud Capron, Isabelle K. Delattre, Frederic Delaruelle, Pierre Wallemacq, and Roger Karel Verbeeck. "A Simultaneous D-Optimal Designed Study for Population Pharmacokinetic Analyses of Mycophenolic Acid and Tacrolimus Early After Renal Transplantation." *The Journal of Clinical Pharmacology* 52, no. 12 (December 2012): 1833–43. https://doi.org/10.1177/0091270011423661.

Nagy, Del, Areej M. Yassin, and Anol Bhattacherjee. "Organizational Adoption of Open Source Software: Barriers and Remedies." *Communications of the ACM* 53, no. 3 (March 2010): 148–51. https://doi.org/10.1145/1666420.1666457.

Neely, Michael N., Michael G. van Guilder, Walter M. Yamada, Alan Schumitzky, and Roger W. Jelliffe. "Accurate Detection of Outliers and Subpopulations With Pmetrics, a Nonparametric and Parametric Pharmacometric Modeling and Simulation Package for R." *Therapeutic Drug Monitoring* 34, no. 4 (August 2012): 467–76. https://doi.org/10.1097/FTD.0b013e31825c4ba6.

Negoescu, Diana M, Eva A Enns, Brooke Swanhorst, Bonnie Baumgartner, James P Campbell, Mark T Osterman, Konstantinos Papamichael, Adam S Cheifetz, and Byron P Vaughn. "Proactive Vs Reactive Therapeutic Drug Monitoring of Infliximab in Crohn's Disease: A Cost-Effectiveness Analysis in a Simulated Cohort." *Inflammatory Bowel Diseases* 26, no. 1 (January 2020): 103–11. https://doi.org/10.1093/ibd/izz113.

Nelder, J. A., and R. Mead. "A Simplex Method for Function Minimization." *The Computer Journal* 7, no. 4 (January 1965): 308–13. https://doi.org/10.1093/comjnl/7.4.308.

Op den Buijsch, Robert A. M., Afke van de Plas, Leo M. L. Stolk, Maarten H. L. Christiaans, Johannes P. van Hooff, Nas A. Undre, Marja P. van Dieijen-Visser, and Otto Bekers. "Evaluation of Limited Sampling Strategies for Tacrolimus." *European Journal of Clinical Pharmacology* 63, no. 11 (November 2007): 1039–44. https://doi.org/10.1007/s00228-007-0354-9.

Owen, Joel S., and Jill Fiedler-Kelly. *Introduction to Population Pharmacokinetic/Pharmacodynamic Analysis with Nonlinear Mixed Effects Models*. John Wiley & Sons, 2014.

Papamichael, Konstantinos, Thomas Van Stappen, Niels Vande Casteele, Ann Gils, Thomas Billiet, Sophie Tops, Karolien Claes, et al. "Infliximab Concentration Thresholds During Induction Therapy Are Associated With Short-term Mucosal Healing in Patients With Ulcerative Colitis." *Clinical Gastroenterology and Hepatology* 14, no. 4 (2016): 543–49. https://doi.org/10.1016/j.cgh.2015.11.014.

Pearce, Joshua M. “A Review of Open Source Ventilators for COVID-19 and Future Pandemics.” *F1000Research* 9 (April 2020): 218. https://doi.org/10.12688/f1000research.22942.2.

Peck, Richard W. “Precision Dosing: An Industry Perspective.” *Clinical Pharmacology & Therapeutics* 109, no. 1 (January 2021): 47–50. https://doi.org/10.1002/cpt.2064.

Pinheiro, José C., and Douglas M. Bates. “Approximations to the Log-Likelihood Function in the Nonlinear Mixed-Effects Model.” *Journal of Computational and Graphical Statistics* 4, no. 1 (1995): 12–35.

Polasek, Thomas M., Craig R. Rayner, Richard W. Peck, Andrew Rowland, Holly Kimko, and Amin Rostami-Hodjegan. “Toward Dynamic Prescribing Information: Codevelopment of Companion Model-Informed Precision Dosing Tools in Drug Development.” *Clinical Pharmacology in Drug Development* 8, no. 4 (May 2019): 418–25. https://doi.org/10.1002/cpdd.638.

Polasek, Thomas M., Sepehr Shakib, and Amin Rostami-Hodjegan. “Precision Dosing in Clinical Medicine: Present and Future.” *Expert Review of Clinical Pharmacology* 11, no. 8 (August 2018): 743–46. https://doi.org/10.1080/17512433.2018.1501271.

Powers, Rob, Maryam Etezadi-Amoli, Edith M. Arnold, Sara Kianian, Irida Mance, Maxsim Gibiansky, Dan Trietsch, et al. “Smartwatch Inertial Sensors Continuously Monitor Real-World Motor Fluctuations in Parkinson’s Disease.” *Science Translational Medicine* 13, no. 579 (February 2021): eabd7865. https://doi.org/10.1126/scitranslmed.abd7865.

Press, Rogier R, Bart A Ploeger, Jan den Hartigh, Tahar van der Straaten, Johannes van Pelt, Meindert Danhof, Johan W de Fijter, and Henk-Jan Guchelaar. “Explaining Variability in Tacrolimus Pharmacokinetics to Optimize Early Exposure in Adult Kidney Transplant Recipients:” *Therapeutic Drug Monitoring* 31, no. 2 (April 2009): 187–97. https://doi.org/10.1097/FTD.0b013e31819c3d6d.

Radawski, Christine A., Tarek A. Hammad, Susan Colilla, Paul Coplan, Kenneth Hornbuckle, Emily Freeman, Meredith Y. Smith, Rachel E. Sobel, Priya Bahri, and Ariel E. Arias. “The Utility of Real-world Evidence for Benefit-risk Assessment, Communication, and Evaluation of Pharmaceuticals: Case Studies.” *Pharmacoepidemiology and Drug Safety* 29, no. 12 (2020): 1532–39.

Ribba, Benjamin, Sherri Dudal, Thierry Lavé, and Richard W. Peck. “Model-informed Artificial Intelligence: Reinforcement Learning for Precision Dosing.” *Clinical Pharmacology & Therapeutics* 107, no. 4 (2020): 853–57.

Rutgeerts, Paul, William J. Sandborn, Brian G. Feagan, Walter Reinisch, Allan Olson, Jewel Johanns, Suzanne Travers, et al. “Infliximab for Induction and Maintenance Therapy for Ulcerative Colitis.” *New England Journal of Medicine* 353, no. 23 (December 2005): 2462–76. https://doi.org/10.1056/NEJMoa050516.

Saito, Jumpei, Kensuke Shoji, Yusuke Oho, Satoshi Aoki, Shotaro Matsumoto, Michiko Yoshida, Hidefumi Nakamura, et al. “Meropenem Pharmacokinetics During Extracorporeal Membrane Oxygenation and Continuous Haemodialysis: A Case Report.” *Journal of Global Antimicrobial Resistance* 22 (September 2020): 651–55. https://doi.org/10.1016/j.jgar.2020.04.029.

Sanathanan, Lilly P., and Carl C. Peck. “The Randomized Concentration-Controlled Trial: An Evaluation of Its Sample Size Efficiency.” *Controlled Clinical Trials* 12, no. 6 (December 1991): 780–94. https://doi.org/10.1016/0197-2456(91)90041-J.

Scheetz, Marc H, Thomas P Lodise, Kevin J Downes, George Drusano, and Michael Neely. “The Case for Precision Dosing: Medical Conservatism Does Not Justify Inaction.” *Journal of Antimicrobial Chemotherapy* 76, no. 7 (June 2021): 1661–65. https://doi.org/10.1093/jac/dkab086.

Sheiner, Lewis B. “Analysis of Pharmacokinetic Data Using Parametric Models. III. Hypothesis Tests and Confidence Intervals.” *Journal of Pharmacokinetics and Biopharmaceutics* 14, no. 5 (October 1986): 539–55. https://doi.org/10.1007/BF01059660.

Sheiner, Lewis B., Stuart Beal, Barr Rosenberg, and Vinay V. Marathe. “Forecasting Individual Pharmacokinetics.” *Clinical Pharmacology & Therapeutics* 26, no. 3 (September 1979): 294–305. https://doi.org/10.1002/cpt1979263294.

Soetaert, Karline, Thomas Petzoldt, and R. Woodrow Setzer. “Solving Differential Equations in *r* : Package **deSolve**.” *Journal of Statistical Software* 33, no. 9 (2010). https://doi.org/10.18637/jss.v033.i09.

Staatz, Christine E, and Susan E Tett. “Clinical Pharmacokinetics and Pharmacodynamics of Tacrolimus in Solid Organ Transplantation:” *Clinical Pharmacokinetics* 43, no. 10 (2004): 623–53. https://doi.org/10.2165/00003088-200443100-00001.

Størset, Elisabet, Anders \AAsberg, Morten Skauby, Michael Neely, Stein Bergan, Sara Bremer, and Karsten Midtvedt. “Improved Tacrolimus Target Concentration Achievement Using Computerized Dosing in Renal Transplant Recipients-a Prospective, Randomized Study.” *Transplantation* 99, no. 10 (2015): 2158–66. https://doi.org/10.1097/TP.0000000000000708.

Størset, Elisabet, Nick Holford, Stefanie Hennig, Troels K. Bergmann, Stein Bergan, Sara Bremer, Anders Åsberg, Karsten Midtvedt, and Christine E. Staatz. “Improved Prediction of Tacrolimus Concentrations Early After Kidney Transplantation Using Theory-Based Pharmacokinetic Modelling: Theory-Based Tacrolimus PK Modelling.” *British Journal of Clinical Pharmacology* 78, no. 3 (September 2014): 509–23. https://doi.org/10.1111/bcp.12361.

Størset, Elisabet, Nick Holford, Karsten Midtvedt, Sara Bremer, Stein Bergan, and Anders Åsberg. "Importance of Hematocrit for a Tacrolimus Target Concentration Strategy." *European Journal of Clinical Pharmacology* 70, no. 1 (January 2014): 65–77. https://doi.org/10.1007/s00228-013-1584-7.

Syversen, Silje Watterdal, Guro Løvik Goll, Kristin Kaasen Jørgensen, Øystein Sandanger, Joseph Sexton, Inge Christoffer Olsen, Johanna Elin Gehin, et al. "Effect of Therapeutic Drug Monitoring Vs Standard Therapy During Infliximab Induction on Disease Remission in Patients With Chronic Immune-Mediated Inflammatory Diseases: A Randomized Clinical Trial." *JAMA* 325, no. 17 (May 2021): 1744. https://doi.org/10.1001/jama.2021.4172.

Tanabe, Kunio, and Masahiko Sagae. "An Exact Cholesky Decomposition and the Generalized Inverse of the Variance–Covariance Matrix of the Multinomial Distribution, with Applications." *Journal of the Royal Statistical Society: Series B (Methodological)* 54, no. 1 (1992): 211–19.

Van der Stighelen, V., P. Burggraeve, L. Geudens, H. Van Pottelbergh, R. Baeten, H. Claessens, E. Cornelis, and L. Pauwels. "ICT-Tools En Gebruik van EMD Door de Huisarts." *Hefbomen En Barrières Bij Het Gebruik van EMD En Elektronische Communicatie Door Huisartsen, Onderzoeksrapport*, 2012.

Vande Casteele, Niels, Marc Ferrante, Gert Van Assche, Vera Ballet, Griet Compernolle, Kristel Van Steen, Steven Simoens, Paul Rutgeerts, Ann Gils, and Séverine Vermeire. "Trough Concentrations of Infliximab Guide Dosing for Patients With Inflammatory Bowel Disease." *Gastroenterology* 148, no. 7 (June 2015): 1320–1329.e3. https://doi.org/10.1053/j.gastro.2015.02.031.

Vande Casteele, Niels, Hans Herfarth, Jeffry Katz, Yngve Falck-Ytter, and Siddharth Singh. "American Gastroenterological Association Institute Technical Review on the Role of Therapeutic Drug Monitoring in the Management of Inflammatory Bowel Diseases." *Gastroenterology* 153, no. 3 (September 2017): 835–857.e6. https://doi.org/10.1053/j.gastro.2017.07.031.

Vande Casteele, N., J. Jeyarajah, V. Jairath, Brian G. Feagan, and William J. Sandborn. "Infliximab Exposure-Response Relationship and Thresholds Associated With Endoscopic Healing in Patients With Ulcerative Colitis." *Clinical Gastroenterology and Hepatology* 17, no. 9 (2019): P1814–1821. https://doi.org/10.1016/j.cgh.2018.10.036.

Vanhove, Thomas, Pieter Annaert, and Dirk R. J. Kuypers. "Clinical Determinants of Calcineurin Inhibitor Disposition: A Mechanistic Review." *Drug Metabolism Reviews* 48, no. 1 (January 2016): 88–112. https://doi.org/10.3109/03602532.2016.1151037.

Vanhove, Thomas, Hanneke Bouwsma, Luuk Hilbrands, Jesse J. Swen, Isabel Spriet, Pieter Annaert, Bart Vanaudenaerde, Geert Verleden, Robin Vos, and Dirk RJ Kuypers. "Determinants of the Magnitude of Interaction Between Tacrolimus and Voriconazole/Posaconazole in Solid Organ Recipients." *American Journal of Transplantation* 17, no. 9 (2017): 2372–80.

Vanhove, Thomas, Mahmoud Hasan, Pieter Annaert, Stefan Oswald, and Dirk R. J. Kuypers. "Pretransplant 4β-Hydroxycholesterol Does Not Predict Tacrolimus Exposure or Dose Requirements During the First Days After Kidney Transplantation." *British Journal of Clinical Pharmacology* 83, no. 11 (November 2017): 2406–15. https://doi.org/10.1111/bcp.13343.

Velickovic-Radovanovic, R., A. Catic-Djordjevic, J. R. Milovanovic, V. Djordjevic, G. Paunovic, and S. M. Jankovic. "Population Pharmacokinetics of Tacrolimus in Kidney Transplant Patients." *Int. Journal of Clinical Pharmacology and Therapeutics* 48, no. 06 (June 2010): 375–82. https://doi.org/10.5414/CPP48375.

Ven, Kris, Jan Verelst, and Herwig Mannaert. "Should You Adopt Open Source Software?" *IEEE Software* 25, no. 3 (May 2008): 54–59. https://doi.org/10.1109/MS.2008.73.

Vermeire, Severine, Erwin Dreesen, Konstantinos Papamichael, and Marla C. Dubinsky. "How, When, and for Whom Should We Perform Therapeutic Drug Monitoring?" *Clinical Gastroenterology and Hepatology*, 2019. https://doi.org/10.1016/j.cgh.2019.09.041.

Vincenti, Flavio, Stephen C. Jensik, Ronald S. Filo, Joshua Miller, and John Pirsch. "A Long-Term Comparison of Tacrolimus (Fk506) and Cyclosporine in Kidney Transplantation: Evidence for Improved Allograft Survival at Five Years1:" *Transplantation* 73, no. 5 (March 2002): 775–82. https://doi.org/10.1097/00007890-200203150-00021.

Wallemacq, Pierre, Victor W. Armstrong, Merce Brunet, Vincent Haufroid, David W. Holt, Atholl Johnston, Dirk Kuypers, et al. "Opportunities to Optimize Tacrolimus Therapy in Solid Organ Transplantation: Report of the European Consensus Conference." *Therapeutic Drug Monitoring* 31, no. 2 (2009): 139–52. https://doi.org/10.1097/FTD.0b013e318198d092.

Wang, Lingshan, Kimberly Maxfield, Daphne Guinn, Rajanikanth Madabushi, Issam Zineh, and Robert N. Schuck. "A Systematic Assessment of US Food and Drug Administration Dosing Recommendations For Drug Development Programs Amenable to Response-Guided Titration." *Clinical Pharmacology & Therapeutics* 109, no. 1 (January 2021): 123–30. https://doi.org/10.1002/cpt.2068.

Wang, Yaning. “Derivation of Various NONMEM Estimation Methods.” *Journal of Pharmacokinetics and Pharmacodynamics* 34, no. 5 (September 2007): 575–93. https://doi.org/10.1007/s10928-007-9060-6.

Wang, Zhigang, and Erwin Dreesen. “Therapeutic Drug Monitoring of Anti-Tumor Necrosis Factor Agents: Lessons Learned and Remaining Issues.” *Current Opinion in Pharmacology* 55 (December 2020): 53–59. https://doi.org/10.1016/j.coph.2020.09.007.

Wilhelm, Scott, Christopher Carter, Mark Lynch, Timothy Lowinger, Jacques Dumas, Roger A. Smith, Brian Schwartz, Ronit Simantov, and Susan Kelley. “Discovery and Development of Sorafenib: A Multikinase Inhibitor for Treating Cancer.” *Nature Reviews Drug Discovery* 5, no. 10 (October 2006): 835–44. https://doi.org/10.1038/nrd2130.

Winter, Lukas, Ruben Pellicer-Guridi, Lionel Broche, Simone A. Winkler, Henning M. Reimann, Haopeng Han, Felix Arndt, et al. “Open Source Medical Devices for Innovation, Education and Global Health: Case Study of Open Source Magnetic Resonance Imaging.” In *Co-Creation: Reshaping Business and Society in the Era of Bottom-up Economics*, edited by Tobias Redlich, Manuel Moritz, and Jens P. Wulfsberg, 147–63. Cham: Springer International Publishing, 2019. https://doi.org/10.1007/978-3-319-97788-1_12.

Woillard, Jean-Baptiste, Jean Debord, Caroline Monchaud, Franck Saint-Marcoux, and Pierre Marquet. “Population Pharmacokinetics and Bayesian Estimators for Refined Dose Adjustment of a New Tacrolimus Formulation in Kidney and Liver Transplant Patients.” *Clinical Pharmacokinetics* 56, no. 12 (December 2017): 1491–98. https://doi.org/10.1007/s40262-017-0533-5.

Woillard, Jean-Baptiste, Marc Labriffe, Jean Debord, and Pierre Marquet. “Tacrolimus Exposure Prediction Using Machine Learning.” *Clinical Pharmacology & Therapeutics* 110, no. 2 (August 2021): 361–69. https://doi.org/10.1002/cpt.2123.

Wright, Daniel F. B., Jennifer H. Martin, and Serge Cremers. “Spotlight Commentary: Model-informed Precision Dosing Must Demonstrate Improved Patient Outcomes.” *British Journal of Clinical Pharmacology* 85, no. 10 (October 2019): 2238–40. https://doi.org/10.1111/bcp.14050.

Yazdani, Afsaneh, and Eric Boerwinkle. “Causal Inference in the Age of Decision Medicine.” *Journal of Data Mining in Genomics & Proteomics* 6, no. 1 (2015).

Zhang, Shimin, Venkateswaran C. Pillai, Sripal Reddy Mada, Steve Strom, and Raman Venkataramanan. “Effect of Voriconazole and Other Azole Antifungal Agents on Cyp3a Activity and Metabolism of Tacrolimus in Human Liver Microsomes.” *Xenobiotica* 42, no. 5 (May 2012): 409–16. https://doi.org/10.3109/00498254.2011.631224.

Zhang, Yifei, Yaning Wang, Mona Khurana, Hari Cheryl Sachs, Hao Zhu, Gilbert J. Burckart, John Alexander, Lynne P. Yao, and Jian Wang. "Exposure–Response Assessment in Pediatric Drug Development Studies Submitted to the US Food and Drug Administration." *Clinical Pharmacology & Therapeutics* 108, no. 1 (July 2020): 90–98. https://doi.org/10.1002/cpt.1809.

Zhu, Ciyou, Richard H. Byrd, Peihuang Lu, and Jorge Nocedal. "Algorithm 778: L-BFGS-B: Fortran Subroutines for Large-Scale Bound-Constrained Optimization." *ACM Transactions on Mathematical Software* 23, no. 4 (December 1997): 550–60. https://doi.org/10.1145/279232.279236.

Zuo, Xiao-cong, Chee M. Ng, Jeffrey S. Barrett, Ai-jing Luo, Bi-kui Zhang, Chen-hui Deng, Lan-yan Xi, et al. "Effects of Cyp3a4 and Cyp3a5 Polymorphisms on Tacrolimus Pharmacokinetics in Chinese Adult Renal Transplant Recipients: A Population Pharmacokinetic Analysis." *Pharmacogenetics and Genomics* 23, no. 5 (May 2013): 251–61. https://doi.org/10.1097/FPC.0b013e32835fcbb6.

Appendix

The following appendices were provided in the digital copy of this PhD thesis. They are omitted from the paper copy for environmental reasons.

- A summary of PhD jury comments, and the changes made in the PhD thesis to address these.
- Supplementary materials to the chapter on infliximab MIPD simulations
- A list of all precision dosing trials identified in clinicaltrials.gov.
- Code listing for statistical analysis of a tacrolimus MIPD trial.

Appendix A Jury comments on January 2022 manuscript

A.1. Introduction

I want to thank all reviewers for their time and effort to review the thesis manuscript. All comments were addressed, and any changes to the manuscript are summarized in the corresponding reply.

A.2. Sebastian Wicha

A.2.1. General Remarks

The thesis submitted by Ruben Faelens deals with the use of pharmacometric approaches to generate precision dosing tools. The first two introductory chapters provide sufficient details to the reader to understand the basic concepts behind pharmacometrics and precision dosing. Chapter 3 clearly defines the objectives of the thesis, and chapter 4 outlines common methods in pharmacometrics and model-informed precision dosing.

The core results of thesis are described in chapters 5 to 9. Thereby, the results described in chapter 5 describe a true innovation, i.e. a versatile R package to facilitate common processes in model-informed precision dosing. The chapter comprehensively describes the development and structure of the R package and includes illustrative examples.

Chapter 6 describes a simulation study the potential value of model-informed precision dosing in infliximab in patients with ulcerative colitis. It was found that model-informed precision dosing using measured drug exposure led to the lowest variability in exposure and associated disease scores, but might require higher overall drug amounts to achieve this aim.

Chapter 7 and 8 describe the development of a population pharmacokinetic model for tacrolimus in de novo kidney transplant patients. Similarly as in chapter 6, a simulation study was utilized to evaluate the potential value of individualized dosing approaches in this cohort. In particular, the study focused on sample size calculations to plan a potential clinical trail being adequately powered to demonstrate the advantages identified in the simulations study in the clinical setting.

Chapter 9 evaluates dosing strategies of tacrolimus under azole exposure which is frequent in the clinical setting to prevent or treat fungal infections early after transplantation. First a population pharmacokinetic model that successfully quantified the impact of azole therapy on the clearance was developed. The model was subsequently used to develop an adjusted dosing regimen which suggests to use approx. one third of the commonly used dose of tacrolimus in presence of azole co-therapy.

Chapter 10 provides a comprehensive discussion of the chapters of the thesis in relation to the previously defined objectives.

Overall, the thesis is very well written and pleasant to read. The results presented advance the knowledge in the field of pharmacometrics and model-informed precision dosing. In particular the developed software packages represents a remarkable piece of work. The

performed in silico clinical trials represent excellent examples on how to use modelling and simulations to plan clinical trials to evaluate the potential benefits of individualized dosing approaches. It is a minor shortcoming of the thesis that some of the figures are not well formatted (e.g. axis labels missing) or not adequately described. Also, some sections are lacking references to back up some statements made with literature. Nonetheless, I recommend to accept the present thesis with minor revisions.

A.2.2. Major comments

Nil

A.2.3. Minor comments

A.2.3.1. Summary

- Spelling mistake: immunosuppressor Spelling mistake corrected.
- The description of the infliximab simulation study sounds a bit like a real clinical study was conducted. Suggest to rephrase to avoid misunderstandings. Rephrased.
- Just be careful with priority claims on "this is the first work establishing a clear roadmap". Not necessary, but can be dangerous as it is difficult to know everything that has been done so far… Added the qualifier "to the best of our knowledge".

A.2.3.2. Introduction: The road to safe and effective drugs

- Remove 'molar' from EC50 description Removed.
- Reference missing to support mechanism of action for kinase inhibitors. Reference added.

A.2.3.3. Introduction: modeling

- The 'modeling' section does not contain any references. References to the appropriate tutorial papers and books were added.
- Figure 2 on PBPK modeling looks blurry Figure aspect ratio adapted to look less blurry.
- Empiric modeling is described as using a limited number of 'compartments'; could also be phrased as a limited number of 'parameters'. Agreed. As an example, the Friberg turnover model uses a large number of compartments, but only a single parameter describing turnover rate.
- Description of VPC contains a distorted phrasing: "every subject individual between-subject variability" Phrasing adapted.

A.2.3.4. Introduction: MIDD

- "killed early" is colloquial language Language adapted.
- Suggest to substitute "psychopharmica" with "antipsychotic drugs". Suggestion applied.
- Suggest to substitute "blood thinners" with "anticoagulants". Suggestion applied.

- The referencing style is unconventional. When authors are referenced in the text, they are not included in the footnote? The referencing style used is the Chicago 17th manual of style shortened footnote format. The style is documented at https://www.chicagomanualofstyle.org/tools_citationguide.html. We have adapted the style to ensure no duplicate (empty) footnotes will be generated upon citing the same article twice on a single page.

A.2.3.5. Introduction: We are not there yet

- Wrong reference used for UHamburg research describing prediction accuracy. Alihodzic et al described the impact of inaccurate sampling. We apologize for this error. Reference adapted.

A.2.3.6. NLME as a foundation for precision dosing

- 'expensive' -> 'Computationally more expensive' Agreed.

- I understand this [the plots in this chapter] is conceptual, but you may want to add units to the axes and parameters. We agree the current plots seem sloppily labelled. The axes for all plots in this chapter were labelled better.

- The description of omitting'2nlog' from the OFV is a bit vague.

To keep the focus of this chapter on numerical methods, this sentence was simplified. An appropriate reference (Pinheiro 1994) was added for readers wanting the in-depth explanation on 2LL and OFV approximations.

- Methods: Precision Dosing section; why is the algorithm for precision dosing given in code instead of as a mathematical equation?

There is a certain grey area between 'equations' and 'algorithms'. In its current form, we agree with the suggestion to format as an equation. Rather than adapting the formatting, we instead adapted the content to describe the EBE-based precision dosing algorithm, as was our original intent.

A.2.3.7. Methods for MIPD

- Figure 8 is missing the units in the axis labels. Even if this is a conceptual figure, imaginary units were added.

- Avoid using the abbreviation 'MPC/MIPD' in the title 'Residual error models and MPC/MIPD' The abbreviation in the title was retained. We agree abbreviations in titles are to be avoided, yet 'model-predictive control for model-informed precision dosing' is too long as a title, and only using 'model-predictive control' could lead to confusion with model-predictive control algorithms in chemical process control.

- Figure 9 is missing X and Y axis. X and Y axis for longitudinal plot (Time / Concentration) and parameter plots (Time / %change from typical value) added.

- Figure 10: What dose was used for which line? Not sure this (the performed dose adaptation) is self-evident… Might need a bit more context.

The caption was adapted to: *Population typical prediction (blue line) and prediction interval (blue area) for ALT, drug concentration, sum of largest tumor diameter (SLD) and clinical utility (CU) at the administered dose of 1mg. The measured ALT, concentration and sum of largest diameter (red points) inform the EBE fit (red line). If 1mg is continued, SLD is predicted to decrease only slowly, ALT will remain low, and CU is poor. Adapting to 3mg (purple line) is predicted to result in increased efficacy (low SLD), acceptable liver toxicity (low ALT), and better overall clinical utility (CU).*

A.2.3.8. Tdmore

- Change 'NwPharm' to 'MwPharm' Changed.

- Figure caption and units are missing for multiple figures

The following captions were added:

> Concentration-time curve of typical value prediction (blue line), EBE fit (red line and 90% confidence interval area) on observed points (grey points), and prediction for recommended dose (green line and 90% confidence interval, dosing target as grey target reticule) as generated by the code sample above.
>
> Predicted concentration-time curves for an EBE fit (line shapes, cf. figure legend) using 0, 1, 2 or 3 observations (points) for 4 virtual patients (separate panels).

This chapter was built in R package vignette style. The code should match 1:1 with the displayed figures. As adding labels would not serve the intended goal of explaining the code, axis units were purposefully omitted.

A.2.3.9. Infliximab

- Dosing should be formulated as '5 mg', not '5mg' (mind the space). Adapted in-text.

- When referring to tables, this is marked as 'Table Table X'. Resolved in automated generation code.

- NA is reported instead of numbers. This unfortunate coding error was resolved.

A.2.3.10. Quantifying MIPD

- Suggest to change the title to be more precise: Rather: Quantifying the impact of MIPD on…. Or?

Agreed. This was changed to 'Quantifying the impact of MIPD on endpoints'

- Figure 19: the colors of facet A are not explained in the caption.

The colors of facet A were corrected to align with the other figures.

- "MPC/MIPD exceeded this [the theoretical] limit": Is this truly correctly stated? In the plot in Figure 19, the dotted line (i.e. EBE) exceeds this limit, not MPC/MIPD…

There is indeed an error in the visual representation; the dotted line is MPC/MIPD. This was corrected in the figure caption.

- Figure 19 and Figure 20 are referred in the text far from where they are shown.

We chose to move all figures close together, so it is easier to find them.

- Faulty reference ‘ref(fig:tacrolimusPaperTte_all_targets)’

Reference was corrected.

A.2.3.11. Tacrolimus MIPD trial simulation

- Change ‘exercise’ to ‘study’ Agreed.
- Reword the study aim to be more neutral: “improves” -> ‘has an impact’, ‘faster time’ -> ‘difference in time’, etc, both in Methods and Results. Agreed.
- Always use spaces between number and unit. This was corrected throughout the text.
- Figures are not always referred or discussed in the main text.
- Figure 25 is not easy to digest, and requires more explanation. Agreed. More details were added.
- In conclusion, change the word ‘exercise’ to ‘study’.

A.2.3.12. Tacrolimus-azole DDI

- No abstract is provided
- References missing throughout the introduction, particularly on the link CNI -> immune response, variation of daily doses from 3 to 30mg, CYP450 DDI impacting efficacy/safety of tacrolimus.
- Some footnotes are empty. This is an unfortunate result of the citation style, with recurring citations resulting in empty footnotes. This was resolved manually in the final imprimatur version.
- Results section, reference to figure missing.
- Suggest to replace ‘underdosing’ to ‘underexposure’ This was rephrased to suggest both occurs; underexposure is a direct result of the failure to proactively increase the dose at azole discontinuation.
- Model parameter tables: Instead of exp(omega)-1, I guess you mean sqrt(exp(omega^2)-1)? Agreed.
- Figure 38 (ecdf plot) is not discussed in the main text. The plot was incorporated in the main text.

A.3. Geert Verbeke

A.3.1. General Remarks

The manuscript is a solid piece of work in which the candidate shows to know the scientific literature very well. The added value of the work is not so much on the development of new methodology but much more on the implementation of existing models and techniques in accessible software routines, the usefulness of which has been illustrated extensively in two case studies. I have no doubt that this work will be applicable in many other contexts in the future. The candidate is first author on two publications (one published, one accepted) in the peer-reviewed literature, and served as co-author on several other publications.

A.3.2. Major comments

- There is a lot of syntax in the body of various chapters. This hampers readability of the text while not really adding to the understanding of the methodology applied. I therefore recommend moving a lot of the syntax to supplementary material sections, except for Chapter 5, which focuses entirely on the development of the tdmore framework.

We understand that code spread throughout the manuscript may not aid in understanding the methodology applied. This code is targeted towards software engineers wanting to implement the described methods. To meet the needs of both target audiences, we have removed five code sections in chapter "Nonlinear mixed effects modeling", and two more in chapter "Methods for MIPD".

- The topic of this thesis is on precision dosing, which is a particular example of precision medicine or personalized medicine. Yet I failed to find a link to the extensive literature on personalized medicine, which often is based on models for causal inference. Why was a different approach followed here, in comparison to the many other contexts where personalized medicine is applied?

Precision medicine or personalized medicine is an umbrella term which covers many topics. Some of these are entirely unrelated even to statistics. Antibody-drug conjugates seek and bind specific (cancer) cell receptors and deliver toxins directly to the cancer cell, minimizing systemic exposure and chemotherapy-related systemic adverse effects. These types of 'personalized medicine' are not discussed in the thesis.

We thank the reviewer for pointing us to the wide body of research on causal inference for personalized medicine. The key difference between personalized medicine and precision dosing is well described by Yazdani and Boerwinkle (2015). "Personalized medicine is the ability to use an individual's genetic make-up and life experiences to diagnose and treat disease. […] If the response to treatment is a unique characteristic of the individual that cannot be predicted a priori, then true personalized medicine has little practical utility in medicine or biomedical research." Precisely this situation is where MIPD should be applied. The technique employs the up-to-now a posteriori exposure and/or response to a treatment to optimize this ongoing treatment.

We do see exciting opportunities for cross-fertilization of both techniques. Classical interpretation would consider MIPD as just another treatment optimization technique with increased benefit-risk ratio, akin to the switch from skin patch to chewing gum for medicinal nicotine delivery. However, trying a treatment lets you learn about a patient. Whereas traditional causal inference models may include covariates such as "unresponsive to treatment A" in source data, the use of MIPD can increase our knowledge of an individual patient in a more fundamental way. Identified individual pharmacokinetic and

pharmacodynamic parameters correlate across formulation, within-class treatment alternatives, or even across mechanisms. Across-treatment MIPD-supported switching has been poorly studied, but could yield promising results. This should reduce a priori uncertainty of individual response for these treatment alternatives, thereby improving precision of personalized medicine predictions.

The manuscript text has been adapted to include these new insights.

- The text contains many abbreviations. A list of the most used ones would help the readability (maybe replacing the less useful lists of tables and figures).

A glossary was added.

- There are still many figures without a label and a caption (see, e.g., p.42, p.71, p.73, p.75, p.76, p.77, p.79).

Labels and captions were added for most figures, even for figures using imaginary compounds.

- Chapter 8 would benefit from more detailed information about the (statistical) methodology applied. The current text does not allow a methodologist with access to the data to fully reproduce the results reported.

Reproducibility is a key aspect of medicine research. A plethora of research indicates severe problems in reproducing results, either because source data is unavailable, or because methods could not be reproduced. Unfortunately, it is difficult to definitively resolve this comment without indication of what aspects are missing from the text.

To ensure reproducibility for Chapter *"Quantifying MIPD: a test-case for tacrolimus"*, the full source code was published online as Supplementary Materials. This included code to reproduce the source dataset from simulation, thereby respecting data confidentiality and GDPR. We have applied this to Chapter 8, adding the full analysis source code in attachment to the manuscript. This should allow reproduction of all figures and analyses. We trust this resolves the concern voiced by the reviewer.

A.3.3. Minor comments

- P.26 typo: 'Chapter Chapter 4' → 'Chapter 4'
- P.54 typo: 'illustrated in 8' → 'illustrated in Figure 8'

These errors in typesetting were corrected in the final version.

A.4. Philippe Jacqmin

A.4.1. General Remarks

The thesis manuscript meets the quantitative and qualitative requirement for a PhD degree. The manuscript is clear and well organized. The objectives and accomplishments are well and fairly described. The challenges, principles, methods, results and discussions are didactically presented without jeopardizing the complexities.

The manuscript find its originality in:

- the development of the MPC/MIPD method addressing individual time-dependent variability.
- the development of the Tdmore software.
- the automatic connection with HER system.
- the in silico evaluation of data analysis methods (EBE vs MPC/MIPD) in the context of TDM.
- the in silico evaluation (infliximab) and optimization (tacrolimus) of MIPD study designs.
- The revised dose adaptation / recommendation of tacrolimus under azole co-medication.

A.4.2. Major comments

A.4.2.1. Positive comment

The EBE method applied to therapeutic drug monitoring was introduced and implemented in software in the 80's. However, if this method can reasonably address the inter-subject variability, without modifications, it cannot appropriately address significant time-dependent intra-individual variability observed in some conditions such as after organ transplantation. The MCP/MIPD method developed during this PhD offers a plausible solution to this issue. It has been evaluated and validated to some extents during this PhD. It provides a new perspective to clinical pharmacologists and clinicians who apply MIPD. Further fine-tuning (e.g. implementation of d-optimality) and validation (e.g. with other drugs) will probably make it standard.

A.4.2.2. Negative comment

None

A.4.3. Minor comments

- The entire manuscript should be double checked for Figure and Table numbers and for legend and references in the text.

The manuscript was double checked and adapted accordingly.

- Some values are missing at several places in the text (e.g. NA in 6.4.3 and 6.4.4).

This error stemmed from generating these numbers directly from simulated data, while the data could not be found in the right path. This error was corrected.

A.5. Minne Casteels

A.5.1. General Remarks

Comprehensive work, also critically indicating the pitfalls of regulations…

A.5.2. Major comments

none

A.5.3. Minor comments

- Add “Abbreviations” section

A Glossary section was added.

- Minor editing in text

Suggested edits in the manuscript text were fully incorporated.

Appendix B Supplementary Materials for Infliximab Chapter

B.1. Figures and Tables

***Table 9**: Parameter estimates of the original and adapted population pharmacokinetic models.*

Parameter	Original model (OFV=2696)		Adapted model (OFV=2784)	
	Estimate	*(RSE) (%)*	*Estimate*	*(RSE) (%)*
Typical values				
Elimination rate constant (/d)				
baseline Mayo endoscopic subscore 1 (/d)	0.0521	-11	0.0422	-14
baseline Mayo endoscopic subscore 2 (/d)	0.0543	-3.6	0.0463	-3.8
baseline Mayo endoscopic subscore 3 (/d)	0.0667	-8.1	0.0570	-5
albumine	-0.8080	-39		
C-reactive proteine	0.0859	-25		
Volume of distribution (L)	6.3400	-3.2	6.9700	-5
fat-free mass	0.5440	-29	0.5170	-39
Corticosteroids[1]	1.3300	-6.9	1.3000	-7.7
Extensive colitis[1]	1.2300	-1.9E-2	1.2500	-6.9
Interindividual variability[2]				
Elimination rate constant (CV%)	29.8000	-7.8	33.4000	-5.7
Volume of distribution (CV%)	26.5000	-20	23.6000	-41
Interoccasion variability				
Elimination rate constant (CV%)	18.7000	-17	6.7000	-100
Residual variability				
Proportional residual error (CV%)	19.2000	-9.4	32.9000	-5.8
Additive residual error (mg/L)	0.3000	FIX	0.3000	FIX

[1] The corticosteroid and extensive colitis effects were modelled as a fold change compared with the reference of no corticosteroids and no extensive colitis, respectively.[2] Inter-individual CV% was calculated as sqrt(exp(omega)-1). CV, coefficient of variation; OFV, objective function value; RSE: relative standard error.

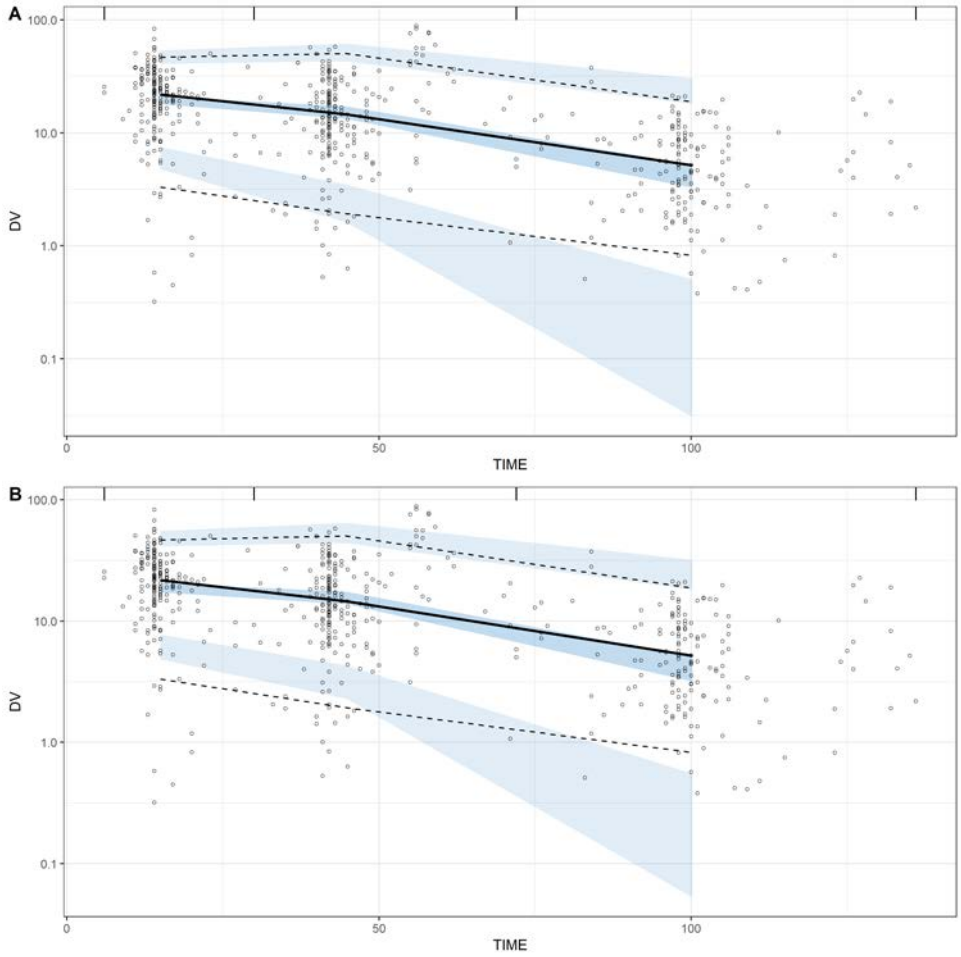


Figure 47*: Prediction-corrected visual predictive check of the original population pharmacokinetic model (a) and the adapted population pharmacokinetic model (b). The solid line connects the observed median prediction-corrected infliximab serum concentrations (mg/L) per bin. The dashed lines connect the 5th and 95th percentiles of the prediction-corrected observations. Shaded areas indicate the 95% confidence interval of the median and 5th and 95th percentiles of the simulated values. The observed prediction-corrected infliximab concentrations are represented by black open circles.*

***Table 10**: Overview of simulation results from sensitivity analysis*

$E_{max,2\rightarrow10}(\%)$		Exposure (mg/L*day)					pEI (%)					Cumulative dose (mg)				
		q05	median	mean	q95	sd	q05	median	mean	q95	sd	q05	median	mean	q95	sd
78%	5 mg/kg	1,049	2,210	2,409	4,448	1,082	38.8	57.2	55.7	67.8	8.96	780	1,080	1,084	1,440	205
	10 mg/kg	2,098	4,419	4,817	8,895	2,165	53.3	66.3	65.1	72.7	6.11	1,560	2,160	2,168	2,880	411
	Covariate-based MIPD	2,302	4,372	4,654	7,940	1,763	54.8	66.1	65.1	72.2	5.46	1,516	2,135	2,151	2,951	431
	Concentration-based MIPD	3,209	4,561	4,666	6,516	1,016	59.8	66.3	66.0	71.2	3.91	1,420	2,213	2,293	3,423	628
80%	5mg/kg	1,044	2,212	2,411	4,468	1,090	38.3	57.3	55.8	68.5	9.31	780	1,080	1,084	1,440	205
	10mg/kg	2,088	4,423	4,821	8,935	2,180	53.3	67.0	65.7	73.9	6.45	1,560	2,160	2,168	2,880	411
	Covariate-based	2,282	4,382	4,650	7,936	1,770	54.8	66.8	65.8	73.3	5.79	1,513	2,133	2,149	2,948	430
	TDM on d14	3,193	4,563	4,657	6,502	1,015	60.1	67.0	66.7	72.2	4.13	1,420	2,210	2,289	3,418	625
100%	5mg/kg	1,052	2,204	2,402	4,418	1,075	34.7	56.1	55.4	73.9	12.00	780	1,080	1,084	1,440	205
	10mg/kg	2,105	4,407	4,805	8,836	2,150	52.0	71.0	70.0	84.7	10.10	1,560	2,160	2,168	2,880	411
	Covariate-based	2,297	4,371	4,641	7,898	1,759	53.9	70.7	70.1	83.5	9.16	1,515	2,133	2,150	2,947	430
	TDM on d14	3,202	4,556	4,657	6,499	1,016	60.9	71.0	71.4	81.4	6.95	1,420	2,211	2,293	3,424	625

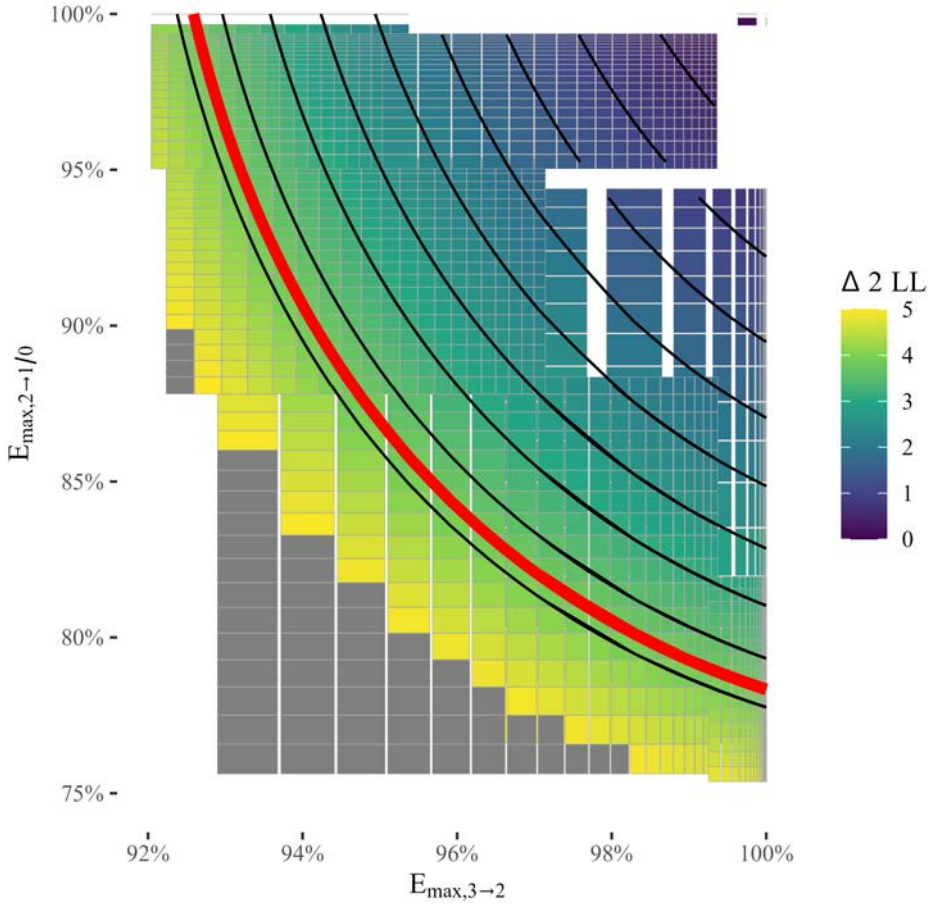


***Figure 48**: Log-likelihood surface for the exposure-response model. The maximum transition probabilities $E_{max,2\rightarrow3}$ and $E_{max,2\rightarrow1/0}$ were fixed while associated AUC50 values (infliximab exposure metrics associated with a half-maximum transition probability) were estimated. The top right point at 100%–100% represents the original exposure-response model as reported by Dreesen et al. All models top right of the red contour line have a Δ2LL below 3.84.*

B.2. NONMEM code of the adapted population pharmacokinetic model

```
$PROBLEM    PopPK analysis of infliximab Induction therapy
```

```
$INPUT      ID TIME DV MDV EVID AMT RATE OCC EXTCOL MPRE CRP
ALB CS WT HT SEX
            ; EXTCOL: Extensive colitis
            ; MPRE:   Mayo endoscopic score pre-induction
            ; CRP:    C-reactive protein
            ; ALB:    Albumin
            ; CS:     Corticosteroids
            ; WT:     Body weight
            ; HT:     Body height
            ; SEX:    Sex (female 0, male 1)

$DATA       PKCOV.csv IGNORE=@ IGNORE(DV .EQ. 0.15) ; exclud
e the BLQ data

$SUBROUTINE ADVAN6 TRANS1 TOL=6

$MODEL      COMP=(CENTRAL,DEFDOS,DEFOBS) COMP=(CUMAUC)

$PK BMI=WT/(HT**2)
    IF (SEX.EQ.1) THEN
                        FFM= (9.27*1000*WT)/((6.68*1000)+(21
6*BMI)) ;male
                  ELSE
                        FFM= (9.27*1000*WT)/((8.78*1000)+(24
4*BMI)) ;female
                  ENDIF

FLAG1=0
FLAG2=0
FLAG3=0
FLAG4=0

IF (OCC .EQ. 1) FLAG1=1
IF (OCC .EQ. 2) FLAG2=1
IF (OCC .EQ. 3) FLAG3=1
IF (OCC .EQ. 4) FLAG4=1

MPRE1=0
MPRE2=0
MPRE3=0
MPRE4=0
```

```
IF (MPRE.EQ.-99) MPRE1=1
IF (MPRE.EQ.1)   MPRE2=1
IF (MPRE.EQ.2)   MPRE3=1
IF (MPRE.EQ.3)   MPRE4=1

TVK = (MPRE1*THETA(4) + MPRE2*THETA(3) + MPRE3*THETA(2) + MP
RE4*THETA(1))
TVV = THETA(6) *(THETA(5)**CS) *((FFM/52)**THETA(7)) *(THETA
(8)**EXTCOL)

K   = TVK * EXP(ETA(1) +FLAG1*ETA(3)+FLAG2*ETA(4)+FLAG3*ETA(
5)+FLAG4*ETA(6))
V   = TVV * EXP(ETA(2))

CL = V*K
S1 = V

$DES DADT(1) =  -K*A(1)
     DADT(2) =  A(1)/S1
     CAUC = A(2) ; cumulative AUC

$ERROR IPRED = F
       IRES  = DV-IPRED
       Y     = IPRED*(1+ERR(1))+ERR(2)
       IWRES = IRES/(SQRT(IPRED**2*SIGMA(1,1)+SIGMA(2,2)))

$THETA (0, 0.057)  ; KE MPRE3
       (0, 0.0463) ; KE MPRE2
       (0, 0.0422) ; KE MPRE1
       (0, 0.617)  ; KE MPRE-99
       (1.3)       ; CS on V
       (6.97)      ; TVV
       (0.517)     ; FFM on V
       (1.25)      ; EXTCOL on V

$OMEGA 0.106.      ;ETA(1)
       0.5         ;ETA(2)
$OMEGA BLOCK(1) 0.054 ;IOV
$OMEGA BLOCK(1) SAME
$OMEGA BLOCK(1) SAME
```

```
$OMEGA BLOCK(1) SAME

$SIGMA 0.103
       0.09 FIX

$ESTIMATION MAXEVAL=9999 NOABORT PRINT=1 NSIG=2 METHOD=1 INT
ER LAPLACE
$COVARIANCE

$TABLE ID DV TIME MDV EVID PRED IPRED RES WRES CWRES IWRES K
CL V ONEHEADER NOPRINT NOAPPEND FILE=PKCOV.sdtab
$TABLE ID K CL V ETA(1) ETA(2) ETA(3) ETA(4) ETA(5) ETA(6) F
IRSTONLY NOAPPEND NOPRINT FILE=PKCOV.patab
$TABLE ID EXTCOL MPRE CS SEX ONEHEADER NOPRINT FILE=PKCOV.ca
tab
$TABLE ID CRP ALB WT HT ONEHEADER NOPRINT FILE=PKCOV.cotab
```

B.3. NONMEM code of the adapted exposure-response model

```
$PROBLEM MPRE==2,3, CAUC drives change

$INPUT ID DV C14 CAUC14 MPRE MPST CDOS CAUC TIPD
       ; TIPD: Time to PD assessment (post-induction endosco
py)

$DATA PKPD.csv IGNORE=@ IGNORE(MPRE.EQ.1) IGNORE(TIPD.LE.30)
IGNORE(TIPD.GE.132)

$PRED PKMETRIC = CAUC
      EMAX1=1/(1+exp(-THETA(6)))
      EMAX2=1/(1+exp(-THETA(7)))

; transition 3 -> 2
X5032  = THETA(1) + ETA(1)
SL32   = THETA(2)
DF32   = (PKMETRIC/X5032)**SL32
D32    = EMAX1*DF32/(1+DF32)

; transition 2 -> 10
X50210 = THETA(3)
```

```
SL210  = THETA(4)
DF210  = (PKMETRIC/X50210)**SL210
D210   = EMAX2*DF210/(1+DF210)

; fraction of patients at 3 and 2 before treatment
P3B    = THETA(5)
P2B    = 1 - THETA(5)

FP33   = (1 - D32)
FP32   = D32 * (1 - D210)
FP310  = D32 * D210

FP23   = (1 - D32)
FP22   = (1 - D210 - (1 - D32))
FP210  = D210

FCHK1  = FP33+FP32+FP310+FP23+FP22+FP210-1 ; -1 compensates
for 2 initial states

; fractions in respective bins
IF (DV.EQ.32)              Y=P3B *FP32    ; 3 -> 2
IF (DV.EQ.31.OR.DV.EQ.30)  Y=P3B *FP310   ; 3 -> <2
IF (DV.EQ.33)              Y=P3B *FP33    ; 3 -> 3
IF (DV.EQ.23)              Y=P2B *FP32    ; 2 -> 3
IF (DV.LT.22)              Y=P2B *FP210   ; 2 -> <2
IF (DV.EQ.22)              Y=P2B *FP22    ; 2 -> 2

FX3   = P3B*FP33 + P2B*FP23
FXN3  = 1-FX3

FX2   = P3B*FP32 + P2B*FP22
FXN2  = 1-FX2

FX10  = P3B*FP310 + P2B*FP210
FXN10 = 1-FX10

$THETA (0, 295, 1000)   ; X5032
       1 FIX            ; SL32
       (0, 1380, 10000) ; X50210
       1 FIX            ; SL210
       (0, 0.516, 1)    ; prob at 3
```

```
       <<REPLACE_DURING_LIKELIHOODPROFILING>> ; EMAX1
       <<REPLACE_DURING_LIKELIHOODPROFILING>> ; EMAX2

$OMEGA 0 FIX

$ESTIMATION MAX=9990 SIGDIG=3 METH=1 LIKE LAPLACE NUMERICAL
NOABORT
$COVARIANCE
```

B.4. tdmore R code for the sensitivity analysis

This file is available online.

B.5. tdmore R code for the simulations

This file is available online.

Appendix C Precision dosing trials in public trial registries

***Table 11**: European and US clinical trials register result for intervention 'Precision dosing OR computer dosing'.*

Trial ID	Start date	Title	Sponsor/Collaborators	Funded by	Publications available?	Protocol available?	Description of stats
NCT00654797	09-2007	Improving Blood Glucose Control With a Computerized Decision Support Tool: Phase 2	Intermountain Health Care, Inc.	Other	no	no	
NCT00655460	02-2006	Improving Blood Glucose Control With a Computerized Decision Support Tool: Phase 1	Intermountain Health Care, Inc.	Other|NIH	no	no	
NCT00733148	07-2004	Correlation Between the Interstitial and Arterial Glucose in Post Surgery Patients	Medical University of Graz	Other	yes	no	
NCT00872079	09-2008	Personalized Warfarin Dosing by Genomics and Computational Intelligence	US Department of Veterans Affairs|VA Office of Research and Development	Other|NIH	no	no	
NCT00993200	08-2009	Personalized Medicine Interface Tool (PerMIT): Warfarin: A Trial Comparing Usual Care Warfarin Initiation to PerMIT Pharmacogenetic Guided Warfarin Therapy	Robert Pendleton|University of Louisville|University of Utah	U.S. Fed	no	no	
NCT01024452	11-2009	Randomized Comparison of Warfarin Dosing Quality Between the Hamilton Nomogram and a Commercial Computer System	Population Health Research Institute	Other	yes	no	
NCT01100723	03-2010	Trial to Optimize Mineral Outcomes in Dialysis Patients	University of Colorado, Denver	Other	yes	yes	a formal power calculation was not performed

Trial ID	Start date	Title	Sponsor/Collaborators	Funded by	Publications available?	Protocol available?	Description of stats
NCT01363193	07-2011	Safety and Efficacy of Lean Body Weight-based IV Heparin Dosing in Obese/Morbidly Obese Patients	Nazareth Hospital	Other	no	no	
NCT01419873	08-2008	Study of a Model-based Approach to Blood Glucose Control in Very-low-birthweight Neonates	Christchurch Women's Hospital	Other	yes	no	
NCT01629251	04-2011	Closing the Loop for 36 Hours in Adolescents With Type 1 Diabetes: Evaluation of Reduced Meal Bolusing	University of Cambridge|Cambridge University Hospitals NHS Foundation Trust	Other|Industry	yes	yes	based on guesstimate
NCT01762059	01-2013	Outpatient Automated Blood Glucose Control With a Bi-hormonal Bionic Endocrine Pancreas	Massachusetts General Hospital|Boston University	Other	yes	no	
NCT01886365	10-2011	Computerized Tight Glycemic Control in Cardiac Surgery	UniversitÃ¤tsklinikum Hamburg-Eppendorf|B. Braun Melsungen AG	Other	yes	no	
NCT01932034	09-2012	Prospective Study to Optimize Vancomycin Dosing in Children and Adults Using Computer Software	Children's Hospital Los Angeles|National Institute of General Medical Sciences (NIGMS)	Other	yes	yes	based on estimate from historic control, and minimally relevant effect size
NCT02010320	01-2014	Computer Guided Doing of Tacrolimus in Renal Transplantation	University of Oslo School of Pharmacy|Rikshospitalet University Hospital	Other	yes	yes	no formal power calculation was performed

Trial ID	Start date	Title	Sponsor/Collaborators	Funded by	Publications available?	Protocol available?	Description of stats
NCT02267408	11-2014	Randomized Controlled Feasibility Trial of the Fearon Algorithm to Improve Management of Unstable Warfarin	McMaster University|Epitome Pharmaceuticals	Other	no	no	
NCT02392364	04-2015	Variable Interval Versus Set Interval Aflibercept for DME	California Retina Consultants|Regeneron Pharmaceuticals	Other	no	no	
NCT02453776	05-2015	Precision Dosing of Infliximab Versus Conventional Dosing of Infliximab	Academisch Medisch Centrum - Universiteit van Amsterdam (AMC-UvA)	Other	yes	yes	intervention group based on estimate from other study; placebo group based on guesstimate
NCT02624037	01-2015	Precision IFX: Using a Dashboard to Individualize Infliximab Dosage	Icahn School of Medicine at Mount Sinai|Prometheus Laboratories	Other	yes	no	
NCT03078491	03-2017	Technological Advances in Glucose Management in Older Adults	Joslin Diabetes Center|Beth Israel Deaconess Medical Center|Boston Children's Hospital|RTI International	Other	yes	no	
NCT03302754	10-2017	Precision Dosing of Alemtuzumab	Children's Hospital Medical Center, Cincinnati	Other|Industry	no	no	
NCT03527238	09-2018	Optimizing Immunosuppression Drug Dosing Via Phenotypic Precision Medicine	University of Florida|National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)	Other	no	no	

Trial ID	Start date	Title	Sponsor/Collaborators	Funded by	Publications available?	Protocol available?	Description of stats
NCT03633656	02-2019	Iron Dosing Pilot Study Using Model Predictive Control	University of Louisville	Other	no	no	
NCT03789591	01-2019	Hydroxyurea Optimization Through Precision Study	Children's Hospital Medical Center, Cincinnati|Doris Duke Charitable Foundation	Other|NIH	yes	yes	based on estimate from other studies
NCT03800875	02-2019	Insulin-plus-pramlintide Closed-loop Strategy to Regulate Glucose Levels Without Carbohydrate Counting	McGill University|Diabetes Canada	Other	yes	yes	based on a guesstimate
NCT03885830	06-2019	Precision Dosing of Tyrosine Kinase Inhibitors in CML Patients	UNC Lineberger Comprehensive Cancer Center	Other	no	no	
NCT03962400	01-2022	Reinforcement Learning for Warfarin Dosing	University of Louisville	Other	no	no	
NCT04340752	01-2021	Optimal Dosing of IC-Green for Visualization of Rotator Cuff Vascularity Using Advanced Imaging Modality Arthroscopy	NYU Langone Health	Other	no	no	
NCT04380311	05-2020	Precision Guided Tacrolimus Dosing in Pediatric Heart Transplant	University of Utah|National Heart, Lung, and Blood Institute (NHLBI)	Other	yes	no	
NCT04666948	12-2020	Precision Dosing of Vancomycin in Critically Ill Children	University Hospital, Ghent|Belgium Health Care Knowledge Centre|Ghent University, Belgium	Other|NIH	no	no	
NCT04822532	06-2021	Precision Dosing of Busulfan in Children Undergoing HSCT	University Hospital, Geneva	Other|Industry	yes	no	

Trial ID	Start date	Title	Sponsor/Collaborators	Funded by	Publications available?	Protocol available?	Description of stats
NCT04911270	12-2021	Clinical Decision Support Tool for Vancomycin Dosing in Children	University of Maryland, Baltimore|Center for Translational Medicine at the School of Pharmacy|Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD)	Other	yes	no	
NCT04974099	10-2021	Personalized Infliximab Induction Strategy With Model-informed Dosing in Patients With Crohn's Disease	Children's Hospital Medical Center, Cincinnati|Crohn's and Colitis Foundation	Other	no	no	
NCT04982172	09-2021	Model-informed Dose De-escalation of Infliximab in Patients With Inflammatory Bowel Diseases	Universitaire Ziekenhuizen Leuven|KU Leuven	Other|Industry	no	no	

Appendix D Tacrolimus MIPD trial simulation: full source code

This chapter contains the full source code for simulation of the proposed clinical trial comparing standard of care to computer-guided dosing for tacrolimus in de novo kidney transplant recipients for the first 14 days following transplant.

D.1. Clinical trial simulation

First, we load the relevant simulation files. We refer to the Supplementary Material of our tacrolimus MIPD simulation (available online at CPT:PSP) for generating these files.

We also define some utility functions for saving and loading results. We augment some simulation results, e.g. adding gaussian blur to clinical results, or adding calculated concentrations for proseval simulation results.

```
## AUthor: Ruben Faelens, 2020-08-15
modelName <- "baseOral_mpc"
TROUGH_DELTA <- 0.01

# setup ------------------------------------------------------
---------------
library(renv)
library(tidyverse)
library(tdmore)
Sys.setenv(PATH=paste(pkgbuild:::rtools_path(),Sys.getenv("P
ATH"), sep=";"))

if(basename(here::here()) == "TacrolimusSimulations") {
  here <- function(...) {
    file.path(here::here(), "SimulationForPaper", ...)
  }
} else {
  here <- here::here
}
```

```
dhere <- function(...) { file.path(here(), modelName, ...)}
ggsave <- function(filename, plot=last_plot(), ...) {
  filename <- dhere(filename)
  ggplot2::ggsave(filename=filename, plot=plot, ...)
  saveRDS(plot, file=paste0(filename, ".RDs"))
}

# Load all files ---------------------------------------------
---------------
scenarioA <- "scenarioA.RDs" %>% dhere() %>% readRDS


physician <- "posthoc.RDs" %>% dhere() %>% readRDS()
physicianConc <- physician %>% unnest(observed) %>%
  mutate(Cwb=Cwb + rnorm(n(), sd=0.001)) %>%  #counter roudi
ng to '12' and '15' exactly by adding gaussian blur
  mutate(Arm="Physician")

proseval <- "proseval.RDs" %>% dhere() %>% readRDS()
prosevalConc <- proseval %>%
  unnest(c(observed, ipred), names_sep=".") %>%
  group_by(ID, OBS) %>%
  mutate(
    INCLUDED=row_number() <= OBS,
    TARGET = (row_number() == OBS[1]+1),
    TARGET2 = (row_number() == OBS[1]+2),
    DAY = floor(observed.TIME / 24),
    DAYFit = ifelse(OBS==0, NA, DAY[OBS]),
    TIME=observed.TIME,
    CwbRatioComputer = ipred.Cwb / observed.Cwb,
    CwbRatioPhysician = 13.5 / observed.Cwb,
    Cwb= 1/CwbRatioComputer * 13.5 #ERROR * TARGET
  )
prosevalConc24h <- prosevalConc %>% filter(DAY-DAYFit == 1)
%>% mutate(Arm="Computer-PE24h")
prosevalConc48h <- prosevalConc %>% filter(DAY-DAYFit == 2)
%>% mutate(Arm="Computer-PE48h")

scenarioAConc <- scenarioA %>%
  group_by(ID) %>%
  filter(row_number() == n()-1) %>% unnest(next_observed) %>
```

```
%  #this is the final result
  select(ID, TIME, Cwb) %>% ungroup() %>%
  mutate(Arm="Computer")

db <- bind_rows(
  physicianConc,
  scenarioAConc
) %>% select(ID, TIME, Cwb, Arm) %>% mutate(
  Arm=factor(Arm),
  DAY = floor(TIME/24)
)
```

We then define a trial simulation framework. A clinical trial result is obtained by bootstrapping from the virtual dataset. A statistical `test` is provided for every trial.

```
# Trial sampling framework ------------------------------
---------------
Ntrial <- 1000
Nsubjects <- 200
NPhys <- round(Nsubjects / 3)
NComp <- round(Nsubjects / 3 * 2)
sampleSet <- db %>% filter(Arm=="Physician") %>%
  group_by(Arm, ID) %>% summarize(n=n()) %>%
  filter(n >= 13) %>% #we want at least 13 of 14 days availa
ble
  pull(ID)

sampleTrial <- function(test, perDay=FALSE, ...) {
  #sample IDs from sampleSet
  IDPhys <- sample(sampleSet, size=NPhys, replace=F)
  IDComp <- sample(sampleSet, size=NComp, replace=F)
  db <- bind_rows(
    db %>% filter(Arm == "Physician" & ID %in% IDPhys) %>% m
utate(ID = ID + 1000),
    db %>% filter(Arm == "Computer" & ID %in% IDComp) %>% mu
tate(ID = ID + 2000)
  )
  if(perDay)
    db <- group_by(db, DAY)
  db %>% do( test(.data, ...) )
}
```

First, an empirical cumulative distribution is provided to understand the simulated data target attainment.

```
# Plot population ECDF ------------------------------------
---------------
scale_fixed <- function(x=scale_color_discrete, values=c(1,
2), n=max(values, na.rm=TRUE), ...) {
  scale <- x()
  discrete_scale(aesthetics=scale$aesthetics,
                 scale_name=scale$scale_name,
                 palette=function(...) {
                   f <- x()$palette
                   val = f(n=n)
                   val[values]
                 },
                 ...
  )
}
methodScales <- list(
  scale_fixed(x=scale_color_discrete, n=3, values=c(1, 3)),
  scale_fixed(x=scale_linetype_discrete, values=c(1, 1)),
  aes(color=Arm, linetype=Arm)
)

ggplot(db, aes(x=DAY, y=Cwb)) +
  geom_hline(yintercept=13.5, linetype=2) +
  annotate("rect", xmin=-Inf, xmax=Inf, ymin=12, ymax=15, fi
ll="green", alpha=0.2) +
  geom_boxplot(aes(group=interaction(Arm,DAY), color=Arm), p
osition=position_dodge(width=0.8), width=0.7) +
  scale_y_log10() +
  labs(x="Time post transplant (days)", y="Concentration (ng
/mL)", color="") +
  theme(legend.position="bottom") +
  methodScales
ggsave("computer_vs_physician_boxplot.png")

last_plot() + geom_text(aes(label=ID, color=Arm), data=. %>%
filter(Arm != "Physician", Cwb > 30))
ggsave("computer_outliers.png")
```

```
q05 <- function(x){quantile(x, 0.05)}
q95 <- function(x){quantile(x, 0.95)}
ggplot(db, aes(x=DAY, y=Cwb, color=Arm, linetype=Arm)) +
  stat_summary(fun=median, fun.min=q05, fun.max=q95,
               position=position_dodge(width=0.6)) +
  stat_summary(geom="errorbar", fun=median, fun.min=q05, fun
.max=q95,
               position=position_dodge(width=0.6), show.lege
nd = F) +
  geom_hline(yintercept=13.5, linetype=2) +
  annotate("rect", xmin=-Inf, xmax=Inf, ymin=12, ymax=15, fi
ll="green", alpha=0.2) +
  #geom_boxplot(aes(group=interaction(Arm,DAY), color=Arm),
position=position_dodge(width=0.8), width=0.7) +
  scale_y_log10() +
  labs(x="Time post transplant (days)", y="Concentration (ng
/mL)", color="") +
  theme(legend.position="bottom") +
  methodScales
ggsave("computer_vs_physician_bars.png")
```

Below, probability of target attainment (PTA) is evaluated. First, the overall simulated PTA in the virtual population is evaluated, after which a clinical trial simulation evaluates PoSS.

```
# Determine population PTA ----------------------------------
------------------
sumstat <- db %>%
  group_by(DAY, Arm) %>%
  summarize(
    n=n(),
    under=sum(Cwb < 12) / n(),
    pta=sum(Cwb > 12 & Cwb < 15)/n(),
    above=sum(Cwb > 15 ) / n()
  )
ptaLimit <- pnorm(c(12-13.5, 15-13.5)/13.5, mean=0, sd=0.187
) %>% diff()
ggplot(sumstat, aes(x=DAY, y=pta)) +
  geom_line(aes(color=Arm, linetype=Arm)) +
  geom_hline(yintercept=ptaLimit) +
  annotate("text", x=-Inf, y=ptaLimit, hjust=-0.1, vjust=-0.
5, label="Theoretical limit") +
```

```
  scale_y_continuous(labels=scales::percent) +
  methodScales
ggsave("computer_vs_physician_pta.png", width=12, height=8)


# Power calculation on PTA ---------------------------------
------------------
library(MESS)

sumstat <- db %>%
  group_by(ID, Arm) %>%
  summarize(
    n=n(),
    under=sum(Cwb < 12) / n(),
    pta=sum(Cwb > 12 & Cwb < 15)/n(),
    above=sum(Cwb > 15 ) / n()
  )

ggplot(sumstat, aes(x=pta)) + geom_histogram() +
  facet_wrap(~Arm)
ptaPopStat <- sumstat %>% group_by(Arm) %>% summarize(m=mean
(pta), s=sd(pta))

# <fct>          <dbl> <dbl>
#   1 Computer 0.379 0.158
# 2 Physician      0.276 0.161

N = seq(10, 500)
pow <- function(N,power, pValue) {
  MESS::power_t_test(n=N/3,
       sd=ptaPopStat$s[ptaPopStat$Arm=="Physician"],
       power=power,
       sig.level = pValue,
       ratio=2,
       sd.ratio=ptaPopStat$s[ptaPopStat$Arm=="Computer"]/pta
PopStat$s[ptaPopStat$Arm=="Physician"],
       #delta=abs(diff(ptaPopStat$m)),
       alternative="one.sided")
}
powerResult <- tidyr::crossing(N=seq(10, 500), powerValue=c(
0.6, 0.8, 0.9), pValue=c(0.01, 0.05)) %>%
```

```
  rowwise() %>%
  mutate(res=
           list(
             unlist(pow(N, powerValue, pValue))
           )
  ) %>% tidyr::unnest_wider(res)
powerResult %>% filter(powerValue==0.8 & pValue == 0.01) %>%
arrange(abs(as.numeric(delta)-0.276*0.33))

ggplot(powerResult, aes(y=as.numeric(delta), x=as.numeric(n1
)+as.numeric(n2))) +
  geom_hline(yintercept=0.276*0.33, linetype=2) +
  geom_line(aes(color=factor(powerValue), linetype=factor(pV
alue))) +
  coord_cartesian(xlim=c(0, 200), ylim=c(0, 0.2)) + geom_hli
ne(yintercept=0.379-0.276) +
  labs(y="Minimum detectable effect", x="Minimum sample size
", color="Power", linetype="Significance level") +
  scale_color_discrete(labels=function(x){paste0(as.numeric(
x)*100, "%")}) +
  scale_y_continuous(labels=function(y){paste0("+", as.numer
ic(y)*100, "%")}) +
  annotate(x=-Inf, y=0.276*0.33, "label", label="Clinically
relevant effect", hjust=0) +
  annotate(x=-Inf, y=0.379-0.276, "label", label="True effec
t", hjust=0)
ggsave("pta_power_mde.png", width=12, height=8)
```

The code below evaluates the time-to-event endpoint. As with PTA, the virtual population is evaluated first, after which PoSS is evaluated.

The `survfit` function creates Kaplan-Meier curves, and the `survdiff()` function is used to compare two survival curves using a Mantel-Haenszel test. There are two 'parameters' to choose when defining the endpoint: minimal number of days in target (we chose ≥ 3) and day on which to evaluate (we chose day 7). Alternatively, overall difference in survival up to day X can be detected; we evaluated day 7, 8, 9 and 10. The key function is `trial.tte.test`.

```
# Population TTE analysis ---------------------------------
-------------------------
library(survival)
```

```
dataSetSurv <- db %>%
  mutate(ID = ID + 100000*as.numeric(Arm) ) %>%
  group_by(ID, Arm) %>%
  mutate(INTARGET = between(Cwb, 12, 15), CUMTARGET=cumsum(I
NTARGET)) %>%
  summarize(TTT = DAY[min(which(CUMTARGET == 3))], #3 days i
n target
            MAXTIME = n()) %>%
  ungroup() %>%
  mutate(SURV = ifelse(is.finite(TTT), 1, 0), #censored, or
event occurred?
         TTT=ifelse(is.finite(TTT), TTT, MAXTIME )) %>%
  as.data.frame()
survdiff(Surv(TTT, SURV) ~ Arm, data=dataSetSurv)
fit1 <- survival::survfit(Surv(TTT, SURV) ~ Arm, data=dataSe
tSurv)
mySum <- summary(fit1)
print(mySum)

broom::tidy(fit1) %>%
  arrange(time) %>%
  filter(time == 7) %>%
  mutate_at(vars(estimate, conf.high, conf.low), ~1-.x)


# Power calculations ----------------------------------------
---------------
## See https://shariq-mohammed.github.io/files/cbsa2019/2-po
wer-and-sample-size.html#3_two-arm_study
hazardValues <- broom::tidy(fit1) %>%
  arrange(time)
# difference in hazard ratio is pPlacebo = pTreatment^DELTA
==>
#     log(pOld) / log(pTreatment) =
TwoArmDeaths = function(Delta, p, alpha, pwr){
  z.alpha = qnorm(alpha, lower.tail=F)
  z.beta = qnorm(1-pwr, lower.tail=F)
  num = (z.alpha + z.beta)^2
  denom = p*(1-p)*(log(Delta))^2
  dd = num/denom
  dd
```

```
}

fun <- function(relReduction, power, alpha) {
  dHR <- log(0.725) / log(1 - (1-0.725)*(1+relReduction))
  nEvents <- TwoArmDeaths(Delta=dHR, p=0.66, alpha=alpha, pw
r=power)
  nAvgEvents <- sum(hazardValues$n.event) / 315 / 2
  #nAvgEvents <- sum(hazardValues$n.event[hazardValues$time
<= 7]) / 315 / 2
  nPatients <- nEvents / nAvgEvents
  nPatients
}
powerResult <- tidyr::crossing(relReduction=seq(0, 1, by=0.0
1), powerValue=c(0.6, 0.8, 0.9), pValue=c(0.01, 0.05)) %>%
  rowwise() %>%
  mutate(N=fun(relReduction, powerValue, pValue))

ggplot(powerResult, aes(y=as.numeric(relReduction), x=N)) +
  geom_hline(yintercept=0.482/0.275 -1, linetype=2) +
  geom_hline(yintercept=0.33) +
  geom_line(aes(color=factor(powerValue), linetype=factor(pV
alue))) +
  coord_cartesian(xlim=c(0, 800), ylim=c(0, 1)) +
  labs(y="Minimum detectable effect", x="Minimum sample size
", color="Power", linetype="Significance level") +
  scale_color_discrete(labels=function(x){paste0(as.numeric(
x)*100, "%")}) +
  scale_y_continuous(labels=function(y){paste0("+", as.numer
ic(y)*100, "%")}) +
  annotate(x=-Inf, y=0.33, "label", label="Clinically releva
nt effect", hjust=0) +
  annotate(x=-Inf, y=0.482/0.275-1, "label", label="True eff
ect", hjust=0)
ggsave("tte_power_mde.png", width=12, height=8)


library(survminer)
z1 <- survminer::ggsurvplot(fit1, data=dataSetSurv, fun=func
tion(x){
  haz <- x
  haz[1] <- 0
```

```
  for(i in seq(2, length(x))) {
    if(x[i] > x[i-1]) {
      haz[i] <- 0
      next
    }
    drop <- x[i-1] - x[i]
    haz[i] <- drop / x[i-1]
  }
  haz
}) %++%
  labs(y="probability of event") %++%
  scale_x_continuous(breaks=seq(0, 15)) %++%
  scale_y_continuous(breaks=seq(0, 1, by=0.2)) %++%
  theme(panel.grid.major = element_line(), panel.grid.minor.
y=element_line(colour="grey87"))
ggsave("tte.marginal.prob.png")

z1 <- survminer::ggsurvplot(fit1,
                            data=dataSetSurv,
                            conf.int = TRUE,
                            pval = TRUE,
                            fun="event",
                            surv.median.line = "hv")
z1 %++%
  scale_x_continuous(breaks=seq(0, 15)) %++%
  scale_y_continuous(breaks=seq(0, 1, by=0.2)) %++%
  theme(panel.grid.major = element_line(), panel.grid.minor.
y=element_line(colour="grey87"))
ggsave("tte_population.png")

positions <- db %>% full_join(
  expand_grid(Arm=unique(db$Arm), DAY=unique(db$DAY), ID=uni
que(db$ID))
) %>% mutate(INTARGET = !is.na(Cwb) & between(Cwb, 12, 15))
%>%
  group_by(ID, Arm) %>% arrange(ID, Arm, DAY) %>%
  mutate(CUMTARGET = cumsum(INTARGET) + 1) %>%
  group_by(Arm, DAY, CUMTARGET) %>% summarize(N = n()) %>%
  bind_rows(
    tibble(DAY=0, Arm=unique(db$Arm), CUMTARGET=1, N=315),
    .,
```

```
    filter(., DAY==14) %>% mutate(DAY = 15) #repeat day 15
  ) %>%
  group_by(Arm, DAY) %>%
  mutate(CUMN = cumsum(N),
      position=1 - CUMN / 315)


ttePlot <- ggplot(positions, aes(x=DAY-0.5, y=position, colo
r=factor(CUMTARGET))) +
  geom_step(aes(linetype=Arm)) +
  ggrepel::geom_label_repel(data=function(x) {
      crossings <- x %>% filter(Arm == "Computer" & position
> 0.5) %>%
        group_by(CUMTARGET) %>% filter(row_number()==1) %>%
select(CUMTARGET, DAY)
      left_join(crossings, x) #show the days where there is
a crossing
    },
                    aes(x=DAY, y=position, label=CUMTARGET,
                                        fontface=ifelse(Arm=="Ph
ysician", 1, 4)), direction = "y",
                    min.segment.length = 0) +
  #geom_point(data=labels, aes(x=DAY+0.5, y=position)) +
  labs(x="Days post transplant", y="Prop of patients with X
concentrations in target", linetype="") +
  scale_x_continuous(breaks=seq(0, 15)) +
  scale_y_continuous(breaks=seq(0, 1, by=0.2), labels=scales
::percent) +
  theme(legend.position="bottom", panel.grid.minor.x = eleme
nt_blank()) +
  scale_fixed(x=scale_color_discrete, n=3, values=rep(1:3, l
ength.out=20), guide=F)
ggsave("tte_all_targets.png", plot=ttePlot)

# TTE analysis: 9 random trials ----------------------------
----------------------------
plots <- lapply(1:9, function(x){ result <- sampleTrial(func
tion(db){
  dataSetSurv <- db %>% mutate(Arm = factor(Arm)) %>%
    group_by(ID, Arm) %>%
    mutate(INTARGET = between(Cwb, 12, 15), CUMTARGET=cumsum
```

```
(INTARGET)) %>%
    summarize(TTT = DAY[min(which(CUMTARGET == 1))],
              MAXTIME = n()) %>%
    mutate(SURV = ifelse(is.finite(TTT), 1, 0), #censored, o
r event occurred?
           TTT=ifelse(is.finite(TTT), TTT, MAXTIME ))
  s1 <- survival::Surv(time=dataSetSurv$TTT, event=dataSetSu
rv$SURV)
  diff1 <- survdiff(s1 ~ Arm, data=dataSetSurv)
  df <- (sum(1 * (diff1$exp > 0))) - 1
  p.value = pchisq(diff1$chisq, df, lower.tail = FALSE)

  fit1 <- survival::survfit(Surv(TTT, SURV) ~ Arm, data=data
SetSurv)
  z1 <- survminer::ggsurvplot(fit1, conf.int = TRUE, pval =
FALSE, fun="event", surv.median.line = "hv") %++%
    annotate("text", x=5, y=0.25, label=paste0("p: ", format
.pval(p.value, 1)) )
  #z1 <- z1 %++%
  #scale_y_continuous(breaks=seq(0, 1, by=0.2)) %++%
  #    theme(panel.grid.major = element_line(), panel.grid.m
inor.y=element_line(colour="grey87"))
  tibble(plot=list(z1))
})
result$plot[[1]]
})
library(gridExtra)
plots_built <- lapply(plots, survminer:::.build_ggsurvplot)
gridExtra::arrangeGrob(grobs=plots_built, nrow=3, ncol=3) %>
%
  ggsave("tte_9trials.png", ., width=16, height=16)

# TTE analysis: study power ------------------------------
---------------
trial.tte.test <- function(db, howMuch=1, cutOff=7) {
  dataSetSurv <- db %>%
    group_by(ID, Arm) %>%
    filter(DAY <= cutOff) %>%
    mutate(INTARGET = between(Cwb, 12, 15), CUMTARGET=cumsum
(INTARGET)) %>%
    summarize(TTT = DAY[min(which(CUMTARGET == howMuch))],
```

```
              MAXTIME = n()) %>%
    mutate(SURV = ifelse(is.finite(TTT), 1, 0), #censored, o
r event occurred?
           TTT=ifelse(is.finite(TTT), TTT, MAXTIME ))
  s1 <- survival::Surv(time=dataSetSurv$TTT, event=dataSetSu
rv$SURV)
  diff1 <- survdiff(s1 ~ Arm, data=dataSetSurv)
  df <- (sum(1 * (diff1$exp > 0))) - 1
  p.value = pchisq(diff1$chisq, df, lower.tail = FALSE)

  ## Calculate difference per day
  fit1 <- survival::survfit(s1 ~ Arm, data=dataSetSurv)
  fit1df <- broom::tidy(fit1) %>% group_by(time) %>% group_m
odify(function(x, ...){
    Physician <- x %>% filter(strata=="Arm=Physician")
    Computer <- x %>% filter(strata=="Arm=Computer")
    p.value = pnorm(Computer$estimate, mean=Physician$estima
te, sd=sqrt(Physician$std.error^2 + Computer$std.error^2))
    tibble(p.value=p.value)
  })
  ### Is the survival on day 3 HIGHER for physician than for
computer
  ### I.e. is the lowest confidence interval for Physician s
till higher than computer median
  ### If so, computer puts patients on target faster on day
3

  tibble(p.value=p.value, p.value.per.day=list(fit1df) )
}

#trial.tte.test(db, howMuch=3)
#debugonce(trial.tte.test)
#sampleTrial(trial.tte.test, howMuch = 3)

p <- progress::progress_bar$new(total=10*1000*4)
trial.tte.result <- tidyr::crossing(trial=1:1000, howMuch=1:
10, cutOff=7:10) %>% group_by(trial, howMuch, cutOff) %>% gr
oup_modify(function(x, keys){
  p$tick()$print()
  sampleTrial(trial.tte.test, howMuch = keys$howMuch, cutOff
=keys$cutOff)
```

```
}) %>% ungroup()
saveRDS(trial.tte.result, dhere("trial.tte.result.RDs"))
trial.tte.result <- readRDS(dhere("trial.tte.result.RDs"))

ggplot(filter(trial.tte.result, howMuch <= 5), aes(x=p.value
)) +
  geom_hline(yintercept=0.8, linetype=2) +
  stat_ecdf(aes(color=factor(howMuch))) +
  coord_cartesian(xlim=c(0, 0.05)) +
  ggrepel::geom_label_repel(data = . %>% group_by(howMuch, c
utOff) %>% summarize(PoSS=mean(p.value<0.02)),
                            aes(x=0.02, y=PoSS, label=howMuc
h, color=factor(howMuch))) +
  labs(x="p-value", y="PoSS") +
  scale_y_continuous(breaks=seq(0, 1, by=0.2), labels=scales
::percent) +
  scale_color_discrete(guide=F) +
  facet_wrap(~cutOff, labeller=function(var) {
    var$cutOff <- paste("Day:",var$cutOff)
    var
  })
ggsave("tte_pvalue.png")

trial.tte.result %>% group_by(howMuch, cutOff) %>% summarize
(
  PoSS_0.01 = mean(p.value < 0.01),
  PoSS_0.05 = mean(p.value < 0.05)
) %>% filter(howMuch==3)

ggplot(trial.tte.result %>% select(-p.value) %>% unnest(p.va
lue.per.day) %>% filter(time==7), aes(x=p.value)) +
  stat_ecdf(aes(color=factor(howMuch))) +
  coord_cartesian(xlim=c(0, 0.05)) +
  ggrepel::geom_label_repel(data = . %>% group_by(howMuch) %
>% summarize(PoSS=mean(p.value<0.005)),
                            aes(x=0.005, y=PoSS, label=howMu
ch, color=factor(howMuch))) +
  labs(x="p-value", y="PoSS") +
  scale_y_continuous(breaks=seq(0, 1, by=0.2)) +
  scale_color_discrete(guide=F)
ggsave("tte_pvalue_on_day7.png")
```

```
trial.tte.result %>% group_by(howMuch) %>% summarize(PoSS_0.
01 = mean(p.value < 0.01), PoSS_0.05 = mean(p.value < 0.05)
)

trial.tte.result %>% dplyr::select(-p.value) %>% unnest(p.va
lue.per.day) %>%
  filter(between(time, 2, 13)) %>%
  ggplot(aes(x=p.value, color=factor(time))) +
  stat_ecdf() + coord_cartesian(xlim=c(0, 0.05)) +
  facet_wrap(~time, labeller=label_both) +
  scale_color_discrete(guide=F) +
  labs(x="p-value", y="PoSS") +
  scale_y_continuous(breaks=seq(0, 1, by=0.2)) +
  facet_wrap(~howMuch)
ggsave("tte_pvalue_per_time.png")
```

Below, a Cox proportional hazards model is used instead. As the dataset does not have a single hazard ratio, this was not retained.

```
# TTE Analysis through coxph survival model ----------------
---------------
dataSetSurv <- db %>% group_by(ID, Arm) %>%
  filter(DAY <= 7) %>%
  mutate(INTARGET = between(Cwb, 12, 15), CUMTARGET=cumsum(I
NTARGET)) %>%
  summarize(TTT = DAY[min(which(CUMTARGET == 1))],
            MAXTIME = n()) %>%
  ungroup() %>%
  mutate(SURV = ifelse(is.finite(TTT), 1, 0), #censored, or
event occurred?
         TTT=ifelse(is.finite(TTT), TTT, MAXTIME ))
fit2 <- coxph(Surv(TTT, SURV) ~ Arm, data=dataSetSurv)

survminer::ggadjustedcurves(fit2, data=as.data.frame(dataSet
Surv), method="conditional", variable="Arm", fun="event")

## Test if coxph fits
summary(fit2)
if( any( summary(fit2)$logtest['pvalue'] > 0.05 ) ) stop("Co
x proportional-hazards model does not fit!")
```

```r
## Test for deviation of proportional-hazards assumption
zph <- cox.zph(fit2)
print(zph)
if(any( zph$table[,"p"]  < 0.05 )) stop("Cox proportional-ha
zards model does not apply!")
ggcoxzph(zph)

ggcoxdiagnostics(fit2, type="schoenfeld", ox.scale="time")
ggsave("coxph.schoenfeld.png")

p.value <- broom::tidy(fit2)$p.value
fit2Fit <- survfit(fit2, newdata=distinct(dataSetSurv, Arm)
%>% as.data.frame)
summary(fit2Fit)
broom::tidy(fit2Fit) %>% tidyr::pivot_longer(estimate.1:conf
.low.2) %>%
  mutate(
    i = as.numeric(stringr::str_extract(name, "\\d+$")),
    name = substr(name, 1, nchar(name)-2),
    Arm = levels(dataSetSurv$Arm)[i],
    strata = paste0("Arm=", Arm)
  ) %>% select(-i) %>%
  pivot_wider()

# TTE Analysis through coxph count model: NOT APPROPRIATE --
------------------------------

## Analysis through a count model is not appropriate
## This assumes counts are fully at random
## Therefore, the number of patients with a count = 2 would
be:
##    hazard_day1 * hazard_day2
## This is not appropriate
dataSetSurv <- db %>%
  mutate(ID = ID + 100000*as.numeric(Arm) ) %>%
  group_by(ID, Arm) %>%
  filter(n() >= 14) %>%
  mutate(INTARGET = between(Cwb, 12, 15)) %>%
  mutate(time=DAY, #starting time for interval
         time2=ifelse(row_number() == n(), DAY+1, lead(DAY))
, #end time for interval
```

```
          event=INTARGET) %>%
  ungroup() %>%
  as.data.frame()
s1 <- with(dataSetSurv, Surv(time, time2, event, type="count
ing", origin=0) )
fit1 <- survfit(Surv(time, time2, event, type="counting", or
igin=0) ~ Arm, data=dataSetSurv)
summary(fit1)
print(fit1)

fit2 <- coxph(Surv(time, time2, event, type="counting") ~ Ar
m, data=dataSetSurv)

foo <- distinct(dataSetSurv, time, Arm) %>% mutate(time2=tim
e+1, event=TRUE)
foo$PRED <- predict(fit2, newdata=foo, type="expected")
print(foo)
ggplot(foo, aes(x=time, y=PRED)) +
  geom_step(aes(color=Arm, linetype="Predicted")) +
  geom_step(data= dataSetSurv %>% group_by(Arm, time) %>% su
mmarize(PRED=mean(event)), aes(color=Arm, linetype="Observed
")) +
  labs(y="Events during that period")

foo <- distinct(dataSetSurv, time, Arm) %>% mutate(time2=tim
e+1, event=TRUE)
foo$PRED <- predict(fit2, newdata=foo, type="expected")
print(foo)
ggplot(foo, aes(x=time, y=PRED)) +
  geom_step(data=. %>% group_by(time) %>% summarize(PRED=mea
n(c(PRED[1])))), aes(linetype="Predicted")) +
  geom_step(data= dataSetSurv %>% group_by(time) %>% summari
ze(PRED=mean(event)), aes(linetype="Observed")) +
  labs(y="Events during that period")

# TTE Coxph: power analysis ------------------------------
---------------
trial.coxph.test <- function(db, howMuch=1) {
  dataSetSurv <- db %>% group_by(ID, Arm) %>%
    filter(DAY <= 7) %>%
    mutate(INTARGET = between(Cwb, 12, 15), CUMTARGET=cumsum
```

```
(INTARGET)) %>%
    summarize(TTT = DAY[min(which(CUMTARGET == howMuch))],
              MAXTIME = n()) %>%
    ungroup() %>%
    mutate(SURV = ifelse(is.finite(TTT), 1, 0), #censored, o
r event occurred?
           TTT=ifelse(is.finite(TTT), TTT, MAXTIME ))

  fit2 <- coxph(Surv(TTT, SURV) ~ Arm, data=dataSetSurv)
  #if( any( summary(fit2)$logtest['pvalue'] > 0.05 ) ) brows
er()# stop("Cox proportional-hazards model does not fit!")
  #if(any( cox.zph(fit2)$table[,"p"]  < 0.05 )) browser() #s
top("Cox proportional-hazards model does not apply!")

  p.value <- tryCatch({
    broom::tidy(fit2)$p.value
  }, error=function(e){1})
  fit2Fit <- survfit(fit2, newdata=distinct(dataSetSurv, Arm
) %>% as.data.frame)

  fitdf <- tryCatch({
    broom::tidy(fit2Fit) %>% tidyr::pivot_longer(estimate.1:
conf.low.2) %>%
      mutate(
        i = as.numeric(stringr::str_extract(name, "\\d+$")),
        name = substr(name, 1, nchar(name)-2),
        Arm = levels(dataSetSurv$Arm)[i],
        strata = paste0("Arm=", Arm)
      ) %>% select(-i) %>%
      pivot_wider()
  }, error=function(e){tibble()})
  fit1df <- fitdf %>% group_by(time) %>% group_modify(functi
on(x, ...){
    Physician <- x %>% filter(strata=="Arm=Physician")
    Computer <- x %>% filter(strata=="Arm=Computer")
    p.value = pnorm(Computer$estimate, mean=Physician$estima
te, sd=sqrt(Physician$std.error^2 + Computer$std.error^2))
    tibble(p.value=p.value)
  })

  tibble(p.value=p.value, p.value.per.day=list(fit1df) )
```

```
}
trial.coxph.test(db)
p <- progress_estimated(1000*10)
trial.coxph.result <- tidyr::crossing(trial=1:1000, howMuch=
1:10) %>% group_by(trial, howMuch) %>% group_modify(function
(x, keys){
  p$tick()$print()
  sampleTrial(trial.coxph.test, howMuch = keys$howMuch)
})
ggplot(trial.coxph.result, aes(x=p.value)) +
  stat_ecdf(aes(color=factor(howMuch))) +
  coord_cartesian(xlim=c(0, 0.05)) +
  ggrepel::geom_label_repel(data = . %>% group_by(howMuch) %
>% summarize(PoSS=mean(p.value<0.005)),
                            aes(x=0.005, y=PoSS, label=howMu
ch, color=factor(howMuch))) +
  labs(x="p-value", y="PoSS") +
  scale_y_continuous(breaks=seq(0, 1, by=0.2)) +
  scale_color_discrete(guide=F)
ggsave("coxph_tte_pvalue.png")

ggplot(trial.coxph.result, aes(x=p.value)) +
  stat_ecdf(aes(color=factor(howMuch), linetype="CoxPH")) +
  stat_ecdf(data=trial.tte.result, aes(color=factor(howMuch)
, linetype="TTE")) +
  coord_cartesian(xlim=c(0, 0.05)) +
  ggrepel::geom_label_repel(data = . %>% group_by(howMuch) %
>% summarize(PoSS=mean(p.value<0.005)),
                            aes(x=0.005, y=PoSS, label=howMu
ch, color=factor(howMuch))) +
  labs(x="p-value", y="PoSS") +
  scale_y_continuous(breaks=seq(0, 1, by=0.2)) +
  labs(linetype="") + theme(legend.position="bottom") +
  scale_color_discrete(guide=F) +
  geom_vline(xintercept=c(0.01, 0.05), linetype=2, alpha=0.2
)
ggsave("coxph_vs_tte_pvalue.png")

trial.coxph.result %>% group_by(howMuch) %>% summarize(PoSS_
0.01 = mean(p.value < 0.01), PoSS_0.05 = mean(p.value < 0.05
) )
```

```
trial.coxph.result %>% dplyr::select(-p.value) %>% unnest(p.
value.per.day) %>%
  filter(between(time, 2, 13)) %>% filter(between(howMuch, 1
, 6)) %>%
  ggplot(aes(x=p.value, color=factor(time))) +
  stat_ecdf() + coord_cartesian(xlim=c(0, 0.05)) +
  facet_wrap(~time, labeller=label_both) +
  scale_color_discrete() +
  labs(x="p-value", y="PoSS") +
  scale_y_continuous(breaks=seq(0, 1, by=0.2)) +
  facet_wrap(~howMuch, labeller=label_both) +
  labs(title="Ability to show difference  in 'X concentratio
n in target'")
ggsave("coxph_tte_pvalue_per_time.png")
```

The below code analyzes whether a markov chain model would be appropriate.

```
# Correlation between attainment yesterday and today -------
---------------
foo <- db %>% mutate(INTARGET = between(Cwb, 12, 15)) %>% gr
oup_by(ID, Arm) %>% mutate(YESTERDAY = lag(INTARGET)) %>%
  filter(!is.na(YESTERDAY))

## IF you were in target, is there a high probability that y
ou are in target today?
library(markovchain)
trial.mc.test <- function(db) {
  sequence <- db %>% mutate(INTARGET=between(Cwb, 12, 15)) %
>% pull(INTARGET)
  sequenceMatr <- createSequenceMatrix(sequence, sanitize =
FALSE)
  mcFitMLE <- markovchainFit(data = sequence)
  wider <- . %>% as.data.frame() %>%
    as_tibble(rownames="from") %>% pivot_longer(-from, names
_to="to") %>%
    mutate_at(vars(from, to), as.logical)
  mcFitMLE$estimate@transitionMatrix  %>% wider
  #mcFitMLE$standardError %>% wider
  #mcFitMLE$lowerEndpointMatrix
  #mcFitMLE$upperEndpointMatrix
```

```
}
result <- db %>% group_by(ID, Arm) %>% filter(n() >= 13) %>%
group_modify(function(x, keys){trial.mc.test(x)} )
result %>% ungroup() %>% mutate_at(vars(Arm),forcats::fct_re
v) %>% ggplot(aes(x=Arm, y=value)) +
  geom_boxplot() +
  facet_grid(to~from, labeller=label_both)
ggsave("transition_probabilities_boxplot.png")

result %>% ungroup() %>% mutate_at(vars(Arm),forcats::fct_re
v) %>% ggplot(aes(x=value)) +
  geom_density(aes(color=Arm)) +
  geom_vline(data=. %>% group_by(Arm, from, to) %>% summariz
e(mean=mean(value)), aes(xintercept=mean, color=Arm)) +
  geom_text(data=. %>% distinct(from, to) %>% mutate(label=c
ase_when(
    from & to ~ "Stay in target",
    from & !to ~ "Drop out of target",
    !from & to ~ "Get in target",
    !from & !to ~ "Stay out of target"
  )), aes(label=label, x=0, y=Inf), hjust=0, vjust=1) +
  facet_grid(to~from) +
  theme(strip.background = element_blank(),
        strip.text = element_blank()) +
  theme(legend.position="bottom") +
  labs(x="transition probability")
ggsave("transition_probabilities_markovChain.png")

foo %>% mutate_all(as.numeric) %>% group_by(YESTERDAY, Arm)
%>% group_modify(function(x, keys){
  tibble(pNo = mean(!x$INTARGET), pYes = mean(x$INTARGET) )

  #test <- cor.test(~ INTARGET + YESTERDAY, data=x)
  #broom::tidy(test)
  #tibble(test=list( test ) )
})
cor(foo$INTARGET, foo$YESTERDAY)
```

The below evaluates target attainment per individual patient.

```
# Target attainment as ratio per patient ------------------
----------------------
```

```
count <- db %>% group_by(ID, Arm) %>%
  mutate(INTARGET = between(Cwb, 12, 15)) %>%
  summarize(INTARGET = sum(INTARGET), RATIO=sum(INTARGET) /
n() )
descdist(count$RATIO[count$Arm=="Physician"])
fitdist(count$RATIO, "beta", method="mme")
fw <- lapply(c("norm", "gamma", "beta", "logis"), function(x
){
  cat(x, "\n")
  fitdist(count$RATIO[count$Arm=="Computer"], x, method="mme
")
} )
fw <- c(fw, lapply(c("norm", "logis"), function(x){
  cat(x, "\n")
  fitdist(count$RATIO[count$Arm=="Computer"], x, method="mle
")
} ))
denscomp(fw)
qqcomp(fw)
cdfcomp(fw)
ppcomp(fw)
gofstat(fw)

# Ratio of trough samples per patient: power analysis ------
------------------------------
trial.t.test <- function(db, includeDb=FALSE) {
  test1db <- db %>% group_by(ID, Arm) %>%
    mutate(INTARGET = between(Cwb, 12, 15)) %>%
    summarize(INTARGET = sum(INTARGET), RATIO=sum(INTARGET)
/ n() )
  test1 <- t.test(RATIO ~ Arm, data=test1db, alternative="gr
eater")
  res <- tibble(p.value=test1$p.value, estimate=diff(test1$e
stimate), lower.ci=test1$conf.int[1], test=list(test1))
  if(includeDb)
    res$db <- list(test1db)
  res
}

z1 <- db %>% group_by(ID, Arm) %>%
  mutate(INTARGET = between(Cwb, 12, 15)) %>%
```

```
  summarize(INTARGET = sum(INTARGET), RATIO=sum(INTARGET) /
n() ) %>%
  ggplot(aes(x=RATIO)) +
  geom_histogram(aes(fill=Arm), bins=14) +
  facet_grid(Arm~.) +
  geom_vline(aes(xintercept=MEAN), data=. %>% group_by(Arm)
%>% summarize(MEAN=mean(RATIO)) ) +
  geom_text(aes(x=MEAN, label=paste0(round(100*MEAN), "%"),
y=0),
            angle=90, hjust=-0.2, vjust=-0.2, size=8, data=.
%>% group_by(Arm) %>% summarize(MEAN=mean(RATIO)) ) +
  labs(x="Ratio of trough samples in target, per patient") +
  scale_x_continuous(breaks=seq(0, 1, by=0.2), labels=scales
::percent) +
  scale_fill_discrete(guide=F) +
  coord_cartesian(xlim=c(0, 1))
test <- trial.t.test(db)
pText <- paste0("p ", format.pval(test$test[[1]]$p.value), "
\n", round(test$estimate*100, 1), "% [", round(100*test$lowe
r.ci,1), "%, Inf]")
z2 <- z1 + geom_text(data=tibble(Arm="Physician"),
                label=pText, x=0, y=Inf, vjust=1, hjust=0 )
ggsave("samples_in_target_per_patient.png", z1)


# 9 trials ---------------------------------------------------
---------------
result <- purrr::map(1:9, function(i){
  test <- sampleTrial(trial.t.test, includeDb=TRUE)

  pText <- paste0("p ", format.pval(test$test[[1]]$p.value),
"\n", round(test$estimate*100, 1), "% [", round(100*test$low
er.ci,1), "%, Inf]")
  z1 %+% test$db + geom_text(data=tibble(Arm="Physician"),
                         label=pText, x=0, y=Inf, vjust=1, hj
ust=0 )
})
lapply(seq_along(result), function(i) {
  ggsave(paste0("samples_in_target_per_patient_", i, ".png")
, result[[i]], width=4, height=3)
})
```

```
#cowplot::plot_grid(plotlist=result, nrow=3)


# 1000 trials ------------------------------------------------
---------------


result <- tibble(trial=1:1000) %>% group_by(trial) %>% group
_modify(function(...){
  sampleTrial(trial.t.test)
})
result %>% ungroup() %>% arrange(lower.ci) %>% mutate(i=row_
number()) %>%
  ggplot(aes(y=1 - i/max(i))) +
  geom_point(aes(x=-estimate, shape="estimate")) +
  geom_point(aes(x=lower.ci, shape="lower 95% CI")) +
  labs(x="Difference in average per-patient target attainmen
t", y="PoSS") +
  geom_vline(xintercept=0) +
  scale_y_continuous(breaks=seq(0, 1, by=0.2), labels=scales
::percent) +
  scale_x_continuous(labels=scales::percent) +
  labs(shape="") + theme(legend.position="bottom")
ggsave("per_patient_increase.png")
saveRDS(result, dhere("per_patient_increase_data.RDs"))
mean(result$p.value < 0.01)
mean(result$p.value < 0.05)

mean( result$estimate )
```

The code below evaluates the proportion of trough samples in target per day.

```
# Proportion of trough samples in target ------------------
---------------
foo <- db %>%
  mutate(Arm = factor(Arm, levels=c("Physician", "Computer")
, labels=c("Physician", "Computer"))) %>%
  group_by(DAY, Arm) %>%
  mutate(INTARGET = between(Cwb, 12, 15)) %>%
  summarize(X=sum(INTARGET), N=n())
foo %>% group_by(DAY) %>%
  group_modify(function(x, ...){
```

```
    test1 <- prop.test(x=x$X, n=x$N, alternative="less")
    x <- x %>% spread(key=Arm, value=X) %>% mutate(Increase=
(Computer-Physician)/N) %>%
      mutate(Increase = paste0( round(Increase*100, 1), "%")
)
    x$p.value = format.pval(test1$p.value, digits=1L)
    x
  }) %>%
  kableExtra::kable() %>%
  kableExtra::save_kable(file = "tmp.html")

rstudioapi::viewer("tmp.html")


# Proportion of trough samples in target: power analysis ---
---------------
trial.prop.test <- function(db) {
  test1db <- db %>% group_by(Arm) %>%
    mutate(INTARGET = between(Cwb, 12, 15)) %>%
    summarize(INTARGET = sum(INTARGET), RATIO=sum(INTARGET)
/ n(), N=n() )
  test1 <- prop.test(x=test1db$INTARGET, n=test1db$N, altern
ative="greater")
  tibble(p.value=test1$p.value, estimate=diff(test1$estimate
), lower.ci=test1$conf.int[1], test=list(test1))
}
result <- tibble(trial=1:1000) %>% group_by(trial) %>% group
_modify(function(...){
  sampleTrial(trial.prop.test, perDay=TRUE)
})
result %>% filter(between(DAY, 2, 13)) %>%
  ggplot(aes(x=p.value)) +
  geom_vline(xintercept=c(0.01, 0.05), linetype=2) +
  stat_ecdf(pad=F) +
  scale_y_continuous(breaks=seq(0, 1, by=0.2)) +
  scale_x_continuous(breaks=seq(0, 0.05, by=0.01), labels=c(
"", 0.01, "", "", "", 0.05), minor_breaks=c()) +
  coord_cartesian(xlim=c(0, 0.05)) +
  facet_wrap(~DAY, labeller=label_both) +
  labs(x="p value", y="PoSS")
ggsave("prop.test.trials.png")
```

```
result %>% filter(between(DAY, 2, 13)) %>%
  group_by(DAY) %>% arrange(lower.ci) %>% mutate(i = row_num
ber()) %>%
  ggplot(aes(x=-estimate, y=i)) +
  geom_point(aes(shape="estimate")) +
  geom_vline(xintercept=0) +
  scale_x_continuous(labels=scales::percent) +
  geom_point(aes(shape="95% lower CI", x=lower.ci)) +
  facet_wrap(~DAY)
ggsave("prop.test.trials.outcomes.png")
```

The code below evaluates the squared deviation endpoint. A KS-test is employed to detect any differences to the overall statistical distribution.

```
# Squared deviation ------------------------------------
---------------
db <- db %>% mutate(SquaredDeviation = case_when(
  Cwb < 12 ~ (log(Cwb)-log(12))^2,
  Cwb > 15 ~ (log(Cwb)-log(15))^2,
  between(Cwb, 12, 15) ~ 0
))

db %>% group_by(Arm) %>% summarize(m = mean(SquaredDeviation
), sd=sd(SquaredDeviation))

db %>%
  filter(between(DAY, 2, 13)) %>%
  ggplot(aes(x=SquaredDeviation)) +
  stat_ecdf(aes(color=Arm, linetype=Arm), pad = FALSE) +
  facet_wrap(~DAY) +
  geom_text(x=0.2, y=0.5, aes(label=paste0("p: ",format.pval
(p.value, digits=1))),
            data=function(x){
              x %>% group_by(DAY) %>% group_modify(function(
x, ...){
                test1 <- ks.test(x$SquaredDeviation[x$Arm=="
Physician"], x$SquaredDeviation[x$Arm != "Physician"])
                tibble(p.value=test1$p.value)
              })
            })+
  coord_cartesian(xlim=c(0, 0.3)) +
```

```
  labs(x="Squared distance to log-transformed target window"
, y="ECDF", color="", linetype="") +
  theme(legend.position="bottom")
ggsave("ks-test-population.png")



# KS-test trial power -------------------------------------
---------------
trial.ks.test <- function(x) {
  test1 <- ks.test(x$SquaredDeviation[x$Arm=="Physician"], x
$SquaredDeviation[x$Arm != "Physician"])
  tibble(p.value=test1$p.value)
}
result <- tibble(trial=1:1000) %>% group_by(trial) %>% group
_modify(function(...){
  sampleTrial(trial.ks.test, perDay=TRUE)
})
result %>% filter(between(DAY, 2, 13)) %>%
  ggplot(aes(x=p.value)) +
  geom_vline(xintercept=c(0.01, 0.05), linetype=2) +
  stat_ecdf(pad=F) +
  scale_y_continuous(breaks=seq(0, 1, by=0.2)) +
  scale_x_continuous(breaks=seq(0, 0.05, by=0.01), labels=c(
"", 0.01, "", "", "", 0.05), minor_breaks=c()) +
  coord_cartesian(xlim=c(0, 0.05)) +
  facet_wrap(~DAY, labeller=label_both) +
  labs(x="p value", y="PoSS")
ggsave("ks.test.trials.png")

result %>% group_by(DAY) %>% summarize(PoSS = mean(p.value <
0.05)) %>% mutate(OK = PoSS > 0.80)
result %>% group_by(DAY) %>% summarize(PoSS = mean(p.value <
0.01)) %>% mutate(OK = PoSS > 0.80)
```

The code below determines the minimum detectable effect size, assuming a normal distribution for PTA.

```
# use simple stats analysis -------------------------------
---------------

# Stats trial power ---------------------------------------
```

```
---------------
library(MESS)
db %>% group_by(Arm, DAY) %>% summarize(m=mean(SquaredDeviat
ion), s=sd(SquaredDeviation))
db %>% group_by(Arm) %>% summarize(m=mean(SquaredDeviation),
s=sd(SquaredDeviation))

ggplot(db, aes(x=SquaredDeviation)) + geom_histogram() +
  facet_wrap(~Arm)
ptaPopStat <- sumstat %>% group_by(Arm) %>% summarize(m=mean
(pta), s=sd(pta))



# <fct>          <dbl> <dbl>
#   1 Computer 0.379 0.158
# 2 Physician      0.276 0.161

N = seq(10, 500)
pow <- function(N,power, pValue) {
  MESS::power_t_test(n=N/3,
                     sd=ptaPopStat$s[ptaPopStat$Arm=="Physic
ian"],
                     power=power,
                     sig.level = pValue,
                     ratio=2,
                     sd.ratio=ptaPopStat$s[ptaPopStat$Arm=="
Computer"]/ptaPopStat$s[ptaPopStat$Arm=="Physician"],
                     #delta=abs(diff(ptaPopStat$m)),
                     alternative="one.sided")
}
powerResult <- tidyr::crossing(N=seq(10, 500), powerValue=c(
0.6, 0.8, 0.9), pValue=c(0.01, 0.05)) %>%
  rowwise() %>%
  mutate(res=
           list(
             unlist(pow(N, powerValue, pValue))
           )
  ) %>% tidyr::unnest_wider(res)
powerResult %>% filter(powerValue==0.8 & pValue == 0.01) %>%
arrange(abs(as.numeric(delta)-0.276*0.33))
```

```
ggplot(powerResult, aes(y=as.numeric(delta), x=as.numeric(n1
)+as.numeric(n2))) +
  geom_hline(yintercept=0.276*0.33, linetype=2) +
  geom_line(aes(color=factor(powerValue), linetype=factor(pV
alue))) +
  coord_cartesian(xlim=c(0, 200), ylim=c(0, 0.2)) + geom_hli
ne(yintercept=0.379-0.276) +
  labs(y="Minimum detectable effect", x="Minimum sample size
", color="Power", linetype="Significance level") +
  scale_color_discrete(labels=function(x){paste0(as.numeric(
x)*100, "%")}) +
  scale_y_continuous(labels=function(y){paste0("+", as.numer
ic(y)*100, "%")}) +
  annotate(x=-Inf, y=0.276*0.33, "label", label="Clinically
relevant effect", hjust=0) +
  annotate(x=-Inf, y=0.379-0.276, "label", label="True effec
t", hjust=0)
ggsave("pta_mde.png")
```

To further analyze/improve the statistical analysis on SquaredDeviation, the use of a `mmrm` model was explored. A boxcox transformation is used to normalize the data.

```
# Fit linear model: power analysis ------------------------
---------------
db <- db %>%
  mutate(Shifted_SquaredDeviation = SquaredDeviation)
#mutate(Shifted_SquaredDeviation = SquaredDeviation + mean(S
quaredDeviation))

ggplot(db, aes(x=Shifted_SquaredDeviation)) + geom_density()
+ coord_cartesian(xlim=c(0, 1))
boxcoxDf <- db %>%
  filter(SquaredDeviation != 0) %>%
  MASS::boxcox(Shifted_SquaredDeviation ~ Arm + DAY + Arm*DA
Y, data=., lambda=seq(-3, 1, length.out=1000))
lambda <- boxcoxDf$x[ which.max(boxcoxDf$y) ]


data <- db %>%
  mutate(T_SquaredDeviation = (Shifted_SquaredDeviation ^ la
```

```
mbda - 1)/lambda) %>%
  filter(DAY != 1)
data %>%
  filter(SquaredDeviation != 0) %>%
  ggplot(aes(sample=T_SquaredDeviation)) +
  geom_qq() +
  geom_qq_line()
ggsave("T_SquaredDeviation_qq.png")
fitdistrplus::fitdist(data$T_SquaredDeviation, "norm") %>%
  fitdistrplus::gofstat()

library(nlme)
data$ID <- data$ID + 10000*as.numeric(data$Arm)
data$ID <- as.numeric( factor(data$ID) )
data$DAY  <- factor(data$DAY)
data$index  <-as.numeric(data$DAY)
fit.cs <- nlme::gls(T_SquaredDeviation ~ Arm * DAY, data = d
ata,
              #observations are correlated between days with
in an individual, but uncorrelated between different individ
uals
              corr = corSymm(form= ~ index | ID),
              # allow a different variance per day
              weights=varIdent(form = ~ 1 | DAY),
              verbose=TRUE,
              control = glsControl(msVerbose=TRUE))
saveRDS(fit.cs, dhere("mmrm_pop.RDs"))
fit.cs <- readRDS(dhere("mmrm_pop.RDs"))

lsmeans <- lsmeans::lsmeans(fit.cs, ~Arm)

lsmeans <- lsmeans::lsmeans(fit.cs, ~Arm | DAY)
lsmeans %>%
  broom::tidy(conf.int=TRUE) %>%
  ggplot(aes(x=factor(as.numeric(DAY)), y=estimate, color=Ar
m, fill=Arm)) +
  geom_pointrange(aes(ymin=conf.low, ymax=conf.high), positi
on=position_dodge(width=0.5))

fit.cs
```

```
relative <- lsmeans %>%
  broom::tidy(conf.int=TRUE)

lsm.diff <- emmeans::contrast( lsmeans , ratios=TRUE)
lsm.diff %>%
  broom::tidy() %>%
  filter(contrast == first(contrast) ) %>%
  ggplot(aes(x=factor(as.numeric(DAY)))) +
  geom_pointrange(aes(y=estimate, ymin=estimate-1.96*std.err
or, ymax=estimate+1.96*std.error)) +
  geom_hline(yintercept=0) +
  labs(x="Day") +
  scale_y_continuous(name="Relative contrast ratio (%)", lab
els=scales::percent)
ggsave("gls.effect.size.png")


# Stats power test ----------------------------------------
--------------
fit.cs2 <- nlme::gls(T_SquaredDeviation ~ Arm + 0,
                     data = data,
                     corr = corSymm(form= ~ index | ID),
                     weights=varIdent(form = ~ 1 | DAY),
                     verbose=TRUE,
                     control = glsControl(msVerbose=TRUE))
lsmeans2 <- lsmeans::lsmeans(fit.cs2, ~Arm)
relative2 <- lsmeans2 %>%
  broom::tidy(conf.int=TRUE)
ggplot(relative2, aes(x=Arm, y=-1*estimate, fill=Arm)) +
  geom_col() +
  geom_errorbar(aes(ymin=-1*conf.low, ymax=-1*conf.high))

relative2 %>% select(-Arm) %>% summarize_all( .funs=function
(x){
  x[1]/x[2]
  })

relative2 %>% select(-statistic, -p.value, -df) %>%
  tidyr::pivot_wider(names_from=Arm, values_from=c(estimate,
conf.low, conf.high)) %>%
  mutate()
```

```
lsm.diff <- pairs( lsmeans )
lsm.diff %>%
  broom::tidy() %>%
  mutate(estimate = -estimate) %>%
  ggplot(aes(x=DAY)) +
  geom_pointrange(aes(y=estimate, ymin=estimate-1.96*std.err
or, ymax=estimate+1.96*std.error)) +
  geom_hline(yintercept=0) +
  labs(x="Day") +
  scale_y_continuous()



# summary(fmOrth.corSym)$tTable
summary(fit.cs2)$tTable

# C <- corMatrix(fmOrth.corSym$modelStruct$corStruct)[[1]]
# sigmaa <- fmOrth.corSym$sigma *
#   coef(fmOrth.corSym$modelStruct$varStruct, unconstrained
= FALSE)['14']
# ra <- seq(1,0.80,length=nrow(C))
# power.mmrm(N=100, Ra = C, ra = ra, sigmaa = sigmaa, power
= 0.80)


Ra <- corMatrix(fit.cs2$modelStruct$corStruct)[[1]]
sigmaa <- fit.cs2$sigma *
  coef(fit.cs2$modelStruct$varStruct, unconstrained = FALSE)
['14']
ra <- seq(1,0.75,length=nrow(Ra))

fun <- function(N, power, sig.level) {
  longpower::power.mmrm(N=N, Ra = Ra, ra = ra, sig.level=sig
.level, sigmaa = sigmaa, power = power, lambda=2, alternativ
e="one.sided")
}
powerResult <- tidyr::crossing(N=seq(10, 800, by=10), powerV
alue=c(0.6, 0.8, 0.9), pValue=c(0.01, 0.05)) %>%
  rowwise() %>%
  mutate(delta=list(unlist(fun(N, powerValue, pValue)))) %>%
```

```
  tidyr::unnest_wider(delta) %>%
  mutate(delta = as.numeric(delta)/3.477445)

ggplot(powerResult, aes(y=as.numeric(delta), x=N)) +
  geom_hline(yintercept=0.5, linetype=2) +
  geom_hline(yintercept=(3.926175-3.477445) / 3.477445) +
  geom_line(aes(color=factor(powerValue), linetype=factor(pV
alue))) +
  coord_cartesian(xlim=c(0, 800), ylim=c(0, 0.6)) +
  labs(y="Minimum detectable effect (relative)", x="Minimum
sample size", color="Power", linetype="Significance level")
+
  scale_color_discrete(labels=function(x){paste0(as.numeric(
x)*100, "%")}) +
  scale_y_continuous(labels=function(y){paste0("+", as.numer
ic(y)*100, "%")}) +
  annotate(x=-Inf, y=0.5, "label", label="Clinically relevan
t effect", hjust=0) +
  annotate(x=-Inf, y=(3.926175-3.477445)/3.477445, "label",
label="True effect", hjust=0) +
  geom_blank()
ggsave("power_mmrm.png", width=12, height=8)
```

This test was also explored for PoSS. Note that the below code takes a very long time to run.

```
library(lsmeans)
trial.gls.test <- function(db, verbose=TRUE) {
  data <- db %>%
    mutate(T_SquaredDeviation = (Shifted_SquaredDeviation ^
lambda - 1)/lambda) %>%
    filter(DAY != 1)
  data$ID <- data$ID + 10000*as.numeric(data$Arm)
  data$ID <- as.numeric( factor(data$ID) )
  data$DAY  <- factor(data$DAY)
  data$index  <-as.numeric(data$DAY)
  fit.cs <- gls(T_SquaredDeviation ~ Arm*DAY, data = data,
                #observations are correlated between days wi
thin an individual, but uncorrelated between different indiv
iduals
                corr = corSymm(form= ~ index | ID),
                # allow a different variance per day
```

```
                weights=varIdent(form = ~ 1 | DAY),
                verbose=verbose,
                control = glsControl(msVerbose=verbose))
  simpleDiff <- pairs( lsmeans(fit.cs, ~Arm, data=data ) )
  lsm.diff <- pairs( lsmeans(fit.cs, ~Arm | DAY, data=data)
)
  tibble(overall=list(broom::tidy(simpleDiff)), per.day=list
(broom::tidy(lsm.diff)) )
}

sampleTrial(trial.gls.test, verbose=FALSE)

library(nlme)
library(lsmeans)
set.seed(1234)
result <- tibble(trial=1:200) %>% group_by(trial) %>% group_
modify(function(x, keys){
  file <- dhere(paste0("gls-",keys$trial, ".RDs"))
  cat("==================", keys$trial, "\n")
  if(!file.exists(file)) {
    foo <- sampleTrial(trial.gls.test, verbose=FALSE)
    saveRDS(foo, file)
  } else {
    foo <- readRDS(file)
  }
  foo
})

saveRDS(result, dhere("resultGLS.RDs"))
result <- readRDS(dhere("resultGLS.RDs"))

result %>% unnest(overall) %>%
  ggplot(aes(x=p.value)) +
  geom_vline(xintercept=c(0.01, 0.05), linetype=2) +
  stat_ecdf(pad=F) +
  scale_y_continuous(breaks=seq(0, 1, by=0.2)) +
  scale_x_continuous(breaks=seq(0, 0.05, by=0.01), labels=c(
"", 0.01, "", "", "", 0.05), minor_breaks=c()) +
  coord_cartesian(xlim=c(0, 0.05)) +
  labs(x="p value", y="PoSS")
ggsave("gls.overall.png")
```

```
result %>% unnest(per.day) %>% mutate_at(vars(DAY), as.numer
ic) %>% filter(between(DAY, 2, 13)) %>%
  ggplot(aes(x=p.value)) +
  geom_vline(xintercept=c(0.01, 0.05), linetype=2) +
  stat_ecdf(pad=F) +
  scale_y_continuous(breaks=seq(0, 1, by=0.2)) +
  scale_x_continuous(breaks=seq(0, 0.05, by=0.01), labels=c(
"", 0.01, "", "", "", 0.05), minor_breaks=c()) +
  coord_cartesian(xlim=c(0, 0.05)) +
  facet_wrap(~DAY, labeller=label_both) +
  labs(x="p value", y="PoSS")
ggsave("gls.per.day.png")

result %>% unnest(per.day) %>% mutate_at(vars(DAY), as.numer
ic) %>% group_by(DAY) %>% summarize(PoSS = mean(p.value < 0.
05)) %>%
  mutate(OK=PoSS > 0.80)

result %>% unnest(per.day) %>% group_by(DAY) %>% summarize(P
oSS = mean(p.value < 0.01)) %>%
  mutate(OK=PoSS > 0.80)

# Fit linear model to this data ---------------------------
--------------

ggplot(data, aes(sample=T_SquaredDeviation)) +
  stat_qq() +
  stat_qq_line()

data$ID <- data$ID + 1000*as.numeric(data$Arm)
data$DAY  <- factor(data$DAY)
data$index  <-as.numeric(data$DAY)
library(nlme)
data %>%
  filter(Arm == "Physician") %>%
  ggplot(aes(x=DAY, y=SquaredDeviation)) +
  geom_boxplot(aes(color=Arm)) +
  geom_smooth(aes(x=as.numeric(DAY))) +
  coord_cartesian(ylim=c(0, 0.2))
fit1 <- gls(SquaredDeviation ~ Arm + DAY, data=data)
```

```
plot(fit1, form=resid(., type = "p") ~ as.numeric(DAY) | Arm
)

fit.cs <- gls(SquaredDeviation ~ Arm + DAY, data = data,
              #observations are correlated between days with
in an individual, but uncorrelated between different individ
uals
              corr = corSymm(form= ~ index | ID),
              # allow a different variance per day
              weights=varIdent(form = ~ 1 | DAY),
              verbose=TRUE,
              control = glsControl(msVerbose=TRUE))
confint(fit.cs)
fit.cs
summary(fit.cs)
plot(fit.cs)
data %>% mutate(IPRED = fitted(fit.cs)) %>%
  ggplot(aes(x=DAY, y=SquaredDeviation)) +
  geom_boxplot() +
  geom_line(aes(y=IPRED, color=Arm, group=ID)) +
  coord_cartesian(ylim=c(0, 0.5))

plot(fit.cs, form=fitted(.) ~ SquaredDeviation | Arm, )

confint(fit.cs)
qqnorm(fit.cs)

confint(fit.cs)
fitted <- data %>% distinct(Arm, DAY, index)
fitted$T_SquaredDeviation <- predict(fit.cs, newdata=fitted)
ggplot(data, aes(x=DAY, y=T_SquaredDeviation, color=Arm))+
  geom_boxplot() +
  #geom_smooth(aes(x=as.numeric(DAY))) +
  geom_point(data=fitted, size=2, shape=3, position=position
_dodge(width=1))



lsm.diff <- lsmeans(fit.cs, pairwise~Arm*DAY, data=data)
lsmstats <- summary(lsm.diff$lsmeans, level=0.9)
ggplot(lsmstats, aes(x=DAY, color=Arm, fill=Arm)) +
```

```
  #geom_pointrange(aes(y=lsmean, ymin=lower.CL, ymax=upper.C
L),
  #                 position=position_dodge(width=0.6)) +
  geom_boxplot(data=data, aes(y=T_SquaredDeviation), fill=NA
) +
  geom_ribbon(aes(x=as.numeric(DAY), ymin=lower.CL, ymax=upp
er.CL), alpha=0.2)+
  geom_line(aes(x=as.numeric(DAY), y=lsmean)) +
  labs(x="MMRM estimate for SquaredDeviation",
       caption="estimate is still box-cox transformed")
ggsave("mmrm_estimated_effect.png")
```

The below code is provided for completeness and reproducibility. It was used to analyze why the computer performs poorly on Day 2. It was also used to generate interactive plots.

```
ggplot(db, aes(x=Cwb)) +
  geom_vline(xintercept=13.5, linetype=2) +
  stat_ecdf(aes(color=Arm)) +
  coord_cartesian(xlim=c(5, 20)) +
  labs(x="Concentration (ng/mL)", y="ECDF") +
  facet_wrap(~DAY)
ggsave("allsim_ecdf.png")

## Why is 24h-prediction too positive for day 2?
## Because the computer KNOWS that it cannot reach the targe
t concentration!
scenarioA %>% group_by(ID) %>% filter(row_number() == 2) %>%
  do({ predict(.data$iterationFit[[1]], regimen=.data$next_r
egimen[[1]], newdata=.data$next_observed[[1]]) }) %>%
  filter(row_number() == 2) %>%
  ggplot(aes(x=Cwb)) + geom_histogram() +
  geom_vline(xintercept=13.5, linetype=2) +
  labs(x="Concentration (ng/mL)") + ggtitle("Day 2 computer-
predicted concentration after adaptation on day 1")
ggsave("scenarioA_day2_computerFit.png")
# Who is too low?? Some cases where the first
scenarioA %>% group_by(ID) %>% filter(row_number() == 2) %>%
  do({ predict(.data$iterationFit[[1]], regimen=.data$next_r
egimen[[1]], newdata=.data$next_observed[[1]]) }) %>%
  filter(row_number() == 2) %>%
  filter(Cwb < 10) %>% View
```

```
db %>% mutate(CONCCat = cut(Cwb, c(0, 5, 12, 15, 20, Inf)) )
%>%
  mutate(CONCCat = forcats::fct_rev(CONCCat)) %>%
  filter(Arm %in% c("Physician", "Computer")) %>% mutate(Arm
= factor(Arm)) %>%
  mutate(Arm = forcats::fct_rev(Arm)) %>%
  ggplot(aes(x=DAY+as.numeric(Arm)/3 - 0.5 )) +
  geom_bar(aes(fill=CONCCat, color=Arm), width=1/3) +
  scale_x_continuous(breaks=seq(0, 14)) +
  scale_color_manual(values=c(NA, "black"))


# Graphical analysis per patient ------------------------
--------------

## Every row is comparison between arms
db %>%
  filter(Arm %in% c("Physician", "Computer")) %>%
  mutate(Arm = factor(as.character(Arm))) %>% #record in num
eric
  #filter(Arm %in% c("Physician")) %>%
  filter(between(DAY, 1, 14)) %>%
  mutate(CONCCat = cut(Cwb, c(0, 5, 12, 15, 20, Inf)) ) %>%
  mutate_at(vars(CONCCat), forcats::fct_rev) %>%
  mutate_at(vars(CONCCat), forcats::fct_explicit_na) %>%
  filter(ID <= 50) %>%
  ggplot(aes(x=factor(DAY), y=ID+as.numeric(Arm)/2-0.5)) +
  geom_tile(aes(fill=CONCCat, alpha=CONCCat, color=Arm), lin
etype=3) +
  geom_hline(aes(yintercept=ID-0.25)) +
  #geom_point(data=. %>% filter(Cycle != max(Cycle)), aes(co
lor=DDelta), position=position_nudge(x=0.5)) +
  scale_fill_manual(values = c("blue", "lightblue", "green",
"yellow", "purple", "grey")) +
  scale_alpha_manual(values = c(0.2, 0.4, 1, 0.4, 0.2, 0.1))
+
  labs(x="Day", y="Subject ID", fill="Tac (ng/mL)",
       color="Arm", alpha="Tac (ng/mL)",
       title="Concentration data for 100 first subjects",
       subtitle="Ordered based on concentration at day 7") +
```

```
  scale_color_manual(values=c(`Computer`="black", `Physician
`=NA), guide=F)

db %>% filter(Arm == "Physician") %>%
  mutate(CONCCat = cut(Cwb, c(0, 5, 12, 15, 20, Inf)) ) %>%
  mutate_at(vars(CONCCat), forcats::fct_explicit_na) %>%
  group_by(DAY, CONCCat) %>%
  summarize(n=n()) %>%
  group_by(DAY) %>% mutate(pta = round(100* n/sum(n), digits
=1)) %>%
  select(-n) %>%
  mutate_all(function(x){ifelse(is.na(x), 0, x)}) %>%
  spread(key=CONCCat, value=pta) %>%
  kableExtra::kable()



kableExtra::save_kable("tmp.html")
rstudioapi::viewer("tmp.html")

### Comparison between arms using facet
db %>%
  filter(Arm %in% c("Physician", "Computer")) %>%
  mutate(Arm = factor(as.character(Arm))) %>% #record in num
eric
  #filter(Arm %in% c("Physician")) %>%
  filter(between(DAY, 1, 14)) %>%
  mutate(CONCCat = cut(Cwb, c(0, 5, 12, 15, 20, Inf)) ) %>%
  mutate_at(vars(CONCCat), forcats::fct_rev) %>%
  mutate_at(vars(CONCCat), forcats::fct_explicit_na) %>%
  filter(ID <= 100) %>%
  ggplot(aes(x=factor(DAY), y=ID)) +
  geom_tile(aes(fill=CONCCat, alpha=CONCCat)) +
  #geom_point(data=. %>% filter(Cycle != max(Cycle)), aes(co
lor=DDelta), position=position_nudge(x=0.5)) +
  scale_fill_manual(values = c("blue", "lightblue", "green",
"yellow", "purple", "grey")) +
  scale_alpha_manual(values = c(0.2, 0.4, 1, 0.4, 0.2, 0.1))
+
  labs(x="Day", y="Subject ID", fill="Tac (ng/mL)",
       color="Arm", alpha="Tac (ng/mL)",
```

```
       title="Concentration data for 100 first subjects",
       subtitle="Ordered based on concentration at day 7") +
  theme_minimal() +
  facet_grid(~Arm)
ggsave("comparison_individual.png")

# For research seminar: plots ------------------------------
---------------
ggplot(db %>% filter(Arm %in% c("Physician", "Computer") & D
AY==2), aes(x=(Cwb-13.5)^2)) +
  stat_ecdf(aes(color=Arm)) +
  coord_cartesian(xlim=c(0, 100))
ggsave("ks.test.315.day2.png")

ks.test(
  db %>% filter(Arm %in% c("Physician") & DAY==2) %>% mutate
(x=(Cwb-13.5)^2) %>% pull(x),
  db %>% filter(Arm %in% c("Computer") & DAY==2) %>% mutate(
x=(Cwb-13.5)^2) %>% pull(x)
)

# Which statistical test should we use? --------------------
---------------

### Search for appropriate lambda parameter per day
boxcoxDf <- db %>%
  mutate(SquaredDeviation = (Cwb-13.5)^2) %>%
  MASS::boxcox(SquaredDeviation ~ Arm + DAY + Arm*DAY, data=
., lambda=seq(0, 0.3, length.out=100))
lambda <- boxcoxDf$x[ which.max(boxcoxDf$y) ]

foo <- db %>%
  mutate(SquaredDeviation = (Cwb-13.5)^2) %>%
  mutate(T_SquaredDeviation = (SquaredDeviation ^ lambda - 1
)/lambda)

dataSetPhysician <- foo %>% filter(Arm == "Physician") %>% g
roup_by(DAY)
dataSetComputer <- foo %>% filter(Arm == "Computer") %>% gro
up_by(DAY)
Ntrial <- 1000
```

```
NPhys <- 67
NComp <- 200-67


### Sample per day, KS-test
test1 <- function(x) {
  x %>% group_by(trial) %>% group_modify(function(x, ...){
    ks1 <- ks.test(x$SquaredDeviation[x$Arm=="Physician"],
                   x$SquaredDeviation[x$Arm=="Computer"],
                   alternative="less")
    tibble(p.value=ks1$p.value)
  })
}
tibble(trial=1:12) %>% rowwise() %>% do({
  data <- bind_rows(
    dataSetPhysician %>% do({ sample_n(.data, size=NPhys) })
,
    dataSetComputer  %>% do({ sample_n(.data, size=NComp) })
  )
  data$trial <- .data$trial
  data
}) %>% filter(DAY == 4) %>%
  ggplot(aes(x=SquaredDeviation)) +
  stat_ecdf(aes(color=Arm)) +
  coord_cartesian(xlim=c(0, 120)) +
  theme_minimal() + theme(legend.position="bottom") +
  geom_point(data=findSegmentDf, aes(x=x0, y=y0)) +
  geom_point(data=findSegmentDf, aes(x=x0, y=y1)) +
  geom_segment(data=findSegmentDf, aes(x=x0, y=y0, xend=x0,
yend=y1)) +
  geom_label(data=test1, aes(x=50, y=0.50, label=paste0("p="
,round(p.value, 3)), fill=p.value<0.05)) +
  scale_fill_manual(values=c(`FALSE`="white", `TRUE`="green"
), guide=F) +
  facet_wrap(~trial, labeller=label_both)  +
  labs(x="(CONC-13.5)^2", y="ECDF")
ggsave("12_trials_ks.test_day4.png", width=16, height=9)
simTrials1 <- tibble(trial=1:Ntrial) %>% rowwise() %>% do({
  data <- bind_rows(
    dataSetPhysician %>% do({ sample_n(.data, size=NPhys) })
,
```

```
    dataSetComputer  %>% do({ sample_n(.data, size=NComp) })
  )
  data$trial <- .data$trial
  data %>% group_by(DAY) %>% do({ test1(.data) })
})

ggplot(simTrials1 %>% filter(between(DAY, 2, 13)), aes(x=p.v
alue)) +
  stat_ecdf(color="red") +
  coord_cartesian(xlim=c(0, 0.05))+
  facet_wrap(~DAY) +
  labs(x="p-value", y="power") +
  scale_y_continuous(labels=scales::percent, breaks=seq(0, 1
, by=0.2)) +
  geom_hline(yintercept=0.8) +
  scale_x_continuous(breaks=seq(0, 0.05, by=0.01)) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  theme(panel.grid.minor = element_blank())
ggsave("ks.test.power.png", width=16, height=9)

# Make two random samples
findSegmentDf <- function(data) {
  data %>% group_by(trial) %>% group_modify(function(x, ...)
{
    findSegment(x$SquaredDeviation, x$Arm)
  })
}
findSegment <- function(value, valueArms) {
  arms <- unique(valueArms)
  sample1 <- value[valueArms==arms[1]]
  sample2 <- value[valueArms==arms[2]]
  cdf1 <- ecdf(sample1)
  cdf2 <- ecdf(sample2)
  minMax <- seq(min(sample1, sample2), max(sample1, sample2)
, length.out=length(sample1))
  x0 <- minMax[which( abs(cdf1(minMax) - cdf2(minMax)) == ma
x(abs(cdf1(minMax) - cdf2(minMax))) )]
  y0 <- cdf1(x0)
  y1 <- cdf2(x0)
  data.frame(x0=x0, y0=y0, y1=y1)
```

```
}


ggplot(dat, aes(x = KSD, group = group, colour = group, line
type=group))+
  stat_ecdf(size=1) +
  xlab("mm") +
  ylab("Cumulitive Distibution") +
  geom_segment(data=findSegment, aes(x = x0, y = y0, xend =
x0, yend = y1),
               linetype = "dashed", color = "red") +
  geom_point(data=findSegment, aes(x = x0 , y= y0), color="r
ed", size=1) +
  geom_point(aes(x = x0 , y= y1), color="red", size=1) +
  ggtitle("K-S Test: Sample 1 / Sample 2")


### Sample per day, t-test
test2 <- function(x) {
  ks1 <- t.test(x$T_SquaredDeviation[x$Arm=="Physician"], x$
T_SquaredDeviation[x$Arm=="Computer"] )
  tibble(p.value=ks1$p.value)
}
simTrials2 <- tibble(trial=1:Ntrial) %>% rowwise() %>% do({
  data <- bind_rows(
    dataSetPhysician %>% do({ sample_n(.data, size=NPhys) })
,
    dataSetComputer  %>% do({ sample_n(.data, size=NComp) })
  )
  data %>% group_by(DAY) %>% do({ test2(.data) })
})
bind_rows(
  simTrials1 %>% mutate(Method="ks.test"),
  simTrials2 %>% mutate(Method="t.test")
) %>%
  ggplot(aes(x=p.value)) +
  geom_hline(yintercept=0.80) +
  stat_ecdf(aes(color=Method)) +
  facet_wrap(~ DAY) +
  coord_cartesian(xlim=c(0, 0.05))
ggsave("proseval-power-ks.test_vs_t.test_pvaluesEcdf.png", w
```

```
idth=16, height=9)

## MMRM power test
## See https://stats.idre.ucla.edu/r/seminars/repeated-measures-analysis-with-r/
## https://www.r-bloggers.com/mixed-models-exercise-2-repeated-measurements/
##
foo <- db %>%
  mutate(SquaredDeviation = (Cwb-13.5)^2) %>%
  mutate(T_SquaredDeviation = (SquaredDeviation ^ lambda - 1)/lambda)

IDSampleSet <- foo %>% filter(Arm == "Physician") %>% group_by(ID) %>% filter(n() >= 13) %>% distinct(ID) %>% unlist
Ntrial <- 10000
NPhys <- 67
NComp <- 200-67
library(nlme)
library(lsmeans)
### First try this on a single trial
PhysSet <- tibble(ID=sample(IDSampleSet, NPhys, replace=FALSE))
CompSet <- tibble(ID=sample(IDSampleSet, NComp, replace=FALSE))
data <- bind_rows(
  PhysSet %>% left_join(foo %>% filter(Arm=="Physician"), by="ID"),
  CompSet %>% left_join(foo %>% filter(Arm=="Computer"), by="ID")
)
data$ID <- data$ID + 1000*as.numeric(data$Arm)
data$DAY  <- factor(data$DAY)
data$index  <-as.numeric(data$DAY)
fit.cs <- gls(T_SquaredDeviation ~ Arm * DAY, data = data,
              corr = corSymm(form= ~ index | ID),
              weights=varIdent(form = ~ 1 | DAY),
              verbose=TRUE,
              control = glsControl(msVerbose=TRUE))
lsm.diff <- lsmeans(fit.cs, pairwise~Arm*DAY, data=data)
lsmstats <- summary(lsm.diff$lsmeans, level=0.9)
```

```
ggplot(lsmstats, aes(x=DAY, color=Arm, fill=Arm)) +
  geom_pointrange(aes(y=lsmean, ymin=lower.CL, ymax=upper.CL
),
                  position=position_dodge(width=0.6)) +
  labs(x="MMRM estimate for SquaredDeviation",
       caption="estimate is still box-cox transformed")
ggsave("mmrm_estimated_effect.png")

mmrm <- function(data) { ## simple MMRM test
  cat(".")
  data$ID <- data$ID + 1000*as.numeric(data$Arm)
  data$DAY  <- factor(data$DAY)
  data$index  <-as.numeric(data$DAY)
  fit.H1 <- gls(T_SquaredDeviation ~ Arm * DAY, data = data,
                corr = corSymm(form= ~ index | ID),
                weights=varIdent(form = ~ 1 | DAY),
                verbose = T,
                control = glsControl(msVerbose=TRUE))
  anova.H1 <- anova(fit.H1)
  p.value <- anova.H1["Arm", "p-value"]
  tibble(p.value=p.value)
}

mmrm <- function(data) {
  cat(".")

  data$ID <- data$ID + 1000*as.numeric(data$Arm)
  data$DAY  <- factor(data$DAY)
  data$index  <-as.numeric(data$DAY)
  cat("Fitting H0 model...\n")
  fit.H0 <- gls(T_SquaredDeviation ~ DAY, data = data,
                corr = corSymm(form= ~ index | ID),
                weights=varIdent(form = ~ 1 | DAY),
                verbose = T,
                control = glsControl(msVerbose=TRUE),
                method="ML") #allows to compare both methods
  cat("Fitting H1 model...\n")
  fit.H1 <- gls(T_SquaredDeviation ~ Arm * DAY, data = data,
                corr = corSymm(form= ~ index | ID),
                weights=varIdent(form = ~ 1 | DAY),
                verbose = T,
```

```
                control = glsControl(msVerbose=TRUE),
                method="ML") #allows to compare both methods
  fit.anova <- anova(fit.H0, fit.H1)
  p.value <- fit.anova["fit.H1", "p-value"]

  #anova.H1 <- anova(fit.H1)
  #p.value <- anova.H1["Arm", "p-value"]
  tibble(p.value=p.value)
}


simTrials3 <- tibble(trial=1:100) %>% rowwise() %>% do({
  PhysSet <- tibble(ID=sample(IDSampleSet, NPhys, replace=FA
LSE))
  CompSet <- tibble(ID=sample(IDSampleSet, NComp, replace=FA
LSE))
  data <- bind_rows(
    PhysSet %>% left_join(foo %>% filter(Arm=="Physician"),
by="ID"),
    CompSet %>% left_join(foo %>% filter(Arm=="Computer"), b
y="ID")
  ) #ID should be unique!
  mmrm(data)
})
saveRDS(simTrials3, dhere("simMmrm.RDs"))
simTrials3 %>% ggplot(aes(x=p.value)) +
  geom_hline(yintercept=0.8) +
  stat_ecdf(color='red') +
  coord_cartesian(xlim=c(0, 0.05))
ggsave("mmrm_power.png")


# Which DTarget should we use? ----------------------------
---------------

dbTarget <- bind_rows(
  db %>% mutate(trans="log", distance="target"),
  db %>% mutate(trans="identity", distance="target"),
  db %>% mutate(trans="log", distance="window"),
  db %>% mutate(trans="identity", distance="window")
) %>%
```

```
  filter(Arm %in% c("Physician", "Computer")) %>%
  group_by(trans, distance) %>%
  mutate(
    target = ifelse(distance=="target", 13.5,
                    ifelse(Cwb > 15, 15, ifelse(Cwb < 12, 12
, Cwb)) #distance to window
    ),
    targetTrans = do.call(trans[[1]], list(target)),
    CwbTrans = do.call(trans[[1]], list(Cwb)),
    SquaredDeviation = (targetTrans - CwbTrans)^2,
    type=paste(trans, distance, sep="/")
  )

ggplot(dbTarget, aes(x=SquaredDeviation)) +
  stat_ecdf(aes(color=Arm)) +
  labs(color="") +
  facet_grid(distance ~ trans, scales="free_x", shrink=TRUE)
ggsave("which_transformation.png")

dbTargetSum <- dbTarget %>% group_by(type, DAY) %>%
  do({
    x <- .data
    ks1 <- ks.test(x$SquaredDeviation[x$Arm=="Physician"],
                   x$SquaredDeviation[x$Arm!="Physician"]
    )
    tibble( p.value = ks1$p.value)
  })

view_kable <- function(x, ...){
  tab <- paste(capture.output(knitr::kable(x, ...)), collaps
e = '\n')
  tf <- tempfile(fileext = ".html")
  writeLines(tab, tf)
  rstudioapi::viewer(tf)
}
dbTarget %>% group_by(type, DAY) %>%
  do({
    x <- .data
    ks1 <- ks.test(x$SquaredDeviation[x$Arm=="Physician"],
                   x$SquaredDeviation[x$Arm!="Physician"]
    )
```

```
    tibble( p.value = ks1$p.value)
  }) %>%
  ungroup() %>%
  spread(key=type, value=p.value) %>%
  view_kable()









# Interactive graphs ------------------------------------
--------------
library(gganimate)
theme_set(theme_minimal())
z1 <- proseval %>%
  filter(ID==1) %>%
  ggplot(aes(x=TIME)) +
  geom_line(data=. %>% ungroup() %>% mutate(ipred=map(fit, ~
predict(.x, newdata=0:400))) %>% unnest(ipred), aes(y=Cwb),
alpha=0.5) +
  geom_point(data=. %>% unnest(ipred), aes(y=Cwb, group=1L))
+
  geom_point(data=. %>% unnest(observed) %>% group_by(ID, OB
S) %>%
               mutate(
                 INCLUDED=row_number() <= OBS,
                 TARGET = (row_number() == OBS[1]+1)
               ), aes(y=Cwb, color=INCLUDED, group=1L)) +
  scale_x_continuous(breaks=seq(0, 14*24, by=24), labels=seq
(0, 14)) +
  theme(legend.position="bottom") +
  theme(panel.grid.minor = element_blank()) +
  labs(x="Time post transplant (days)", y="Tac concentration
(ng/mL)", color="Included in fit") +
  transition_states(
    OBS,
    transition_length = 1,
    state_length = 3
```

```
  ) +
  enter_fade() +
  exit_shrink() +
  ease_aes('sine-in-out') +
  ggtitle("Iteration #{closest_state}")

z1 %>%
  animate(duration = 30, fps=5, width=480, height=480)
anim_save("subject-1.gif")

z2 <- proseval %>%
  filter(ID==1) %>%
  ungroup() %>%
  mutate(data=pmap(list(object, ipred, observed), ~residuals
(..1, ..2, ..3, weighted=TRUE))) %>%
  unnest(data) %>%
  mutate(lastObs = TIME[ifelse(OBS==0, NA, OBS)] ) %>%
  ggplot(aes(x=TIME, y=Cwb)) +
  #geom_vline(aes(xintercept=lastObs)) +
  geom_point(aes(color=TIME <= lastObs, group=1L)) +
  geom_smooth(se=F) +
  scale_x_continuous(breaks=seq(0, 14*24, by=24), labels=seq
(0, 14)) +
  theme(legend.position="bottom") +
  theme(panel.grid.minor = element_blank()) +
  labs(x="Time post transplant (days)", y="Prediction error
IWRES", color="Included in fit") +
  transition_states(
    OBS,
    transition_length = 1,
    state_length = 2
  ) +
  enter_fade() +
  exit_shrink() +
  ease_aes('sine-in-out') +
  ggtitle("Iteration #{closest_state}")
z2 %>%
  animate(duration = 30, fps=5)
anim_save("subject-1-IWRES.gif")
```

```
gg_color_hue <- function(n) {
  hues = seq(15, 375, length = n + 1)
  hcl(h = hues, l = 65, c = 100)[1:n]
}


proseval %>%
  filter(ID==1) %>%
  unnest(c(ipred,observed), names_sep=".") %>% group_by(ID,
OBS) %>%
  mutate(
    INCLUDED=row_number() <= OBS,
    TARGET = (row_number() == OBS[1]+1)
  ) %>% filter(TARGET) %>%
  ggplot(aes(x=ipred.TIME)) +
  geom_point(aes(y=ipred.Cwb, color="Predicted")) +
  geom_point(aes(y=observed.Cwb, color="Observed")) +
  labs(color="") +
  scale_x_continuous(breaks=seq(0, 14*24, by=24), labels=seq
(0, 14)) +
  theme(legend.position="bottom") +
  theme(panel.grid.minor = element_blank()) +
  labs(x="Time post transplant (days)", y="Tac concentration
(ng/mL)") +
  scale_color_manual(values=c(gg_color_hue(2)[1], "black"))
ggsave("subject_1_pe24h_conc.png")

proseval %>%
  filter(ID==1) %>%
  unnest(c(ipred,observed), names_sep=".") %>% group_by(ID,
OBS) %>%
  mutate(
    INCLUDED=row_number() <= OBS,
    TARGET = (row_number() == OBS[1]+1)
  ) %>% filter(TARGET) %>%
  ggplot(aes(x=ipred.TIME)) +
  annotate("rect", fill="green", xmin=-Inf, xmax=Inf, ymin=1
2, ymax=15, alpha=0.1) +
  geom_point(aes(y=observed.Cwb / ipred.Cwb * 13.5, color="C
omputer")) +
  labs(color="") +
```

```
  scale_x_continuous(breaks=seq(0, 14*24, by=24), labels=seq
(0, 14)) +
  scale_y_continuous(breaks=c(5, 10, 12, 15, 20)) +
  theme(legend.position="bottom") +
  theme(panel.grid.minor = element_blank()) +
  labs(x="Time post transplant (days)", y="Concentration (ng
/mL)")
ggsave("subject_1_pe24h_conc_peTarget_1.png")
last_plot() + geom_point(aes(y=observed.Cwb, color="Physicia
n"))
ggsave("subject_1_pe24h_conc_peTarget_2.png")


proseval %>%
  unnest(c(ipred,observed), names_sep=".") %>% group_by(ID,
OBS) %>%
  mutate(
    INCLUDED=row_number() <= OBS,
    TARGET = (row_number() == OBS[1]+1)
  ) %>% filter(TARGET) %>%
  mutate(Computer=observed.Cwb / ipred.Cwb * 13.5, Physician
=observed.Cwb) %>%
  gather(key=Arm, value=Cwb, Computer, Physician) %>%
  mutate(DAY = floor(ipred.TIME/24)) %>% filter(DAY <= 12) %
>%
  ggplot(aes(x=ipred.TIME)) +
  annotate("rect", ymin=-Inf, ymax=Inf, xmin=12, xmax=15, fi
ll="green", alpha=0.2) +
  stat_ecdf(aes(x=Cwb, color=Arm)) +
  # geom_rug(data=.  %>% filter(MDV==0) %>% group_by(DAY) %>
% do(enframe(quantile(.data$CONC, probs=c(0.05, 0.50, 0.95))
)) %>%
  # mutate(q = case_when(name == "5%" ~ 0.05,
  #                      name == "50%" ~  0.50,
  #                      name == "95%" ~  0.95)
  # ), aes(x=value, linetype=name), length=unit(0.1, "npc"))
+
  coord_cartesian(xlim=c(5, 20)) +
  labs(y="ECDF(x)", x="tac whole-blood concentration (ng/mL)
") +
  facet_wrap(~DAY, labeller=label_both) +
```

```
  scale_linetype_manual(values=c(2, 1, 2)) +
  theme(legend.position="bottom") + labs(linetype="Quantile"
, color="")
ggsave("proseval_ecdf.png")




db %>% filter(DAY <= 12) %>%
  ggplot(aes(x=TIME)) +
  annotate("rect", ymin=-Inf, ymax=Inf, xmin=12, xmax=15, fi
ll="green", alpha=0.2) +
  stat_ecdf(aes(x=Cwb, color=Arm)) +
  # geom_rug(data=.  %>% filter(MDV==0) %>% group_by(DAY) %>
% do(enframe(quantile(.data$CONC, probs=c(0.05, 0.50, 0.95))
)) %>%
  # mutate(q = case_when(name == "5%" ~ 0.05,
  #                      name == "50%" ~  0.50,
  #                      name == "95%" ~  0.95)
  # ), aes(x=value, linetype=name), length=unit(0.1, "npc"))
+
  coord_cartesian(xlim=c(5, 20)) +
  geom_text(data=. %>% group_by(DAY, Arm) %>% summarize(pta=
mean(between(Cwb, 12, 15))) %>%
              mutate(ptaString = paste0(round(pta*100, 0), "
%")),
            aes(x=5, label=ptaString, y=0.1, color=Arm), pos
ition=position_fill(), hjust=0, size=4) +
  labs(y="ECDF(x)", x="tac whole-blood concentration (ng/mL)
") +
  facet_wrap(~DAY, labeller=label_both) +
  scale_linetype_manual(values=c(2, 1, 2)) +
  theme(legend.position="bottom") + labs(linetype="Quantile"
, color="")
ggsave("proseval_ecdf_all.png")

last_plot() %+% (db %>% filter(DAY <= 12 & Arm != "Physician
"))
ggsave("proseval_ecdf_computers.png")
last_plot() %+% (db %>% filter(DAY <= 12 & Arm %in% c("Physi
cian", "Computer")))
```

```
ggsave("proseval_ecdf_adapt_v_physician.png")




# doseChanges ------------------------------------------------
---------------
db2 <- ("../../../../../Tacrolimus/") %>%
  file.path("3.Analysis Data/PostTransplant_VanHove2017BJCP_
KWS.Rds") %>%
  readRDS()  %>% mutate_if(is.numeric, unclass) #remove have
n-related issues
manualEcdf <- db2 %>% filter(EVID==0 & MDV==0) %>%
  group_by(DAY) %>%
  arrange(CONC) %>% mutate(i=1:n())
db2Doses <- db2 %>% filter(EVID!=0 & MDVReason==0) %>%
  mutate(Cycle=floor((TATx+12) / 24 ) ) %>%  # Cycle 1 is fr
om Day0 12:00 to Day1 12:00. The next cycle is determined ba
sed on measure Day1 08:00.
  full_join(
    tidyr::crossing(Cycle=as.numeric(0:15)) # fill in the ga
ps
  ) %>%
  group_by(ID, Cycle) %>% summarize(DDose=sum(AMT, na.rm=TRU
E)) %>%
  group_by(ID) %>% mutate(
    DDoseTomorrow = lead(DDose),
    DDelta = ifelse(DDoseTomorrow==0 & DDose==0, 1.0, DDoseT
omorrow / DDose)
  ) %>%
  full_join(manualEcdf, by=c("ID", Cycle="DAY")) %>%
  group_by(ID) %>% mutate( CONCTomorrow = lead(CONC) )
compDoses <- scenarioA %>% group_by(ID) %>% filter(row_numbe
r() == n()) %>% unnest(regimen) %>%
  mutate(Cycle=floor((TIME+12) / 24 ) ) %>%  # Cycle 1 is fr
om Day0 12:00 to Day1 12:00. The next cycle is determined ba
sed on measure Day1 08:00.
  full_join(
    tidyr::crossing(Cycle=as.numeric(0:15)) # fill in the ga
ps
  ) %>%
```

```
  group_by(ID, Cycle) %>% summarize(DDose=sum(AMT, na.rm=TRU
E)) %>%
  group_by(ID) %>% mutate(
    DDoseTomorrow = lead(DDose),
    DDelta = ifelse(DDoseTomorrow==0 & DDose==0, 1.0, DDoseT
omorrow / DDose)
  ) %>%
  full_join(db %>% filter(Arm=="Computer"), by=c("ID", Cycle
="DAY")) %>%
  mutate(CONC=Cwb) %>%
  group_by(ID) %>% mutate( CONCTomorrow = lead(CONC) )


scale_color_breaks <- function(..., breaks=c(), colours=c())
{
  values = breaks
  w = max(breaks) - min(breaks)
  expand = 0.10 * w

  scale_color_gradientn(colours=colours, breaks=breaks, resc
aler=function(x, from, to) {
    y <- cut(x, c(-Inf, breaks, Inf) )
    width <- 1 / ( length(breaks) + 1 ) # width of a single
bin
    z <- (as.numeric(y)-1) * width + width/2 #pick the middl
e of the bin
    z
  }, limits=c(min(breaks)-expand, max(breaks)+expand ), oob=
scales::squish, ...)
}

compDoses %>%
  filter(between(Cycle, 2, 13)) %>%
  filter(!is.na(DDelta)) %>%
  ggplot(aes(x=CONC, y=DDelta)) +
  annotate("rect", xmin=12, xmax=15, ymin=-Inf, ymax=Inf, fi
ll="green", alpha=0.2) +
  geom_hline(yintercept=1, linetype=2) +
  geom_vline(xintercept=c(5, 12, 15, 20), linetype=2) +
  #geom_point(aes(color=CONCTomorrow, size=ifelse(between(CO
NCTomorrow, 5, 20) | is.na(CONCTomorrow), 1, 2))) +
```

```
  #scale_size_identity() +
  geom_point() +
  #geom_smooth(data=. %>% filter(), se=F) +
  scale_y_continuous(breaks=c(0.33, 0.5, 1, 2, 3), labels=c(
"/3", "/2", "no change", "x2", "x3"), minor_breaks=c()) +
  coord_trans(y="log10", limx=c(3, 30), limy=c(0.2, 3)) +
  scale_x_continuous(breaks=c(5, 12, 15, 20, 30)) +
  stat_function(fun = function(CONC) {(13.5 / CONC)}, color=
"grey", size=2) +
  #coord_cartesian(xlim=c(3, 25), ylim=c(0.2, 5)) + #, clip=
"off"
  scale_color_breaks(breaks=c(5, 12, 15, 20), colours=c("blu
e", "cyan", "grey", "orange", "red"), na.value="pink" ) +
  #annotate(geom="text", x=5, y=1, label="decrease", hjust=1
.5, vjust=-1, color="grey", angle=90, size=4) +
  #annotate(geom="text", x=5, y=1, label="increase", hjust=-
0.5, vjust=-1, color="grey", angle=90, size=4) +
  theme_minimal() +
  #geom_smooth(data=db2Doses, se=F) +
  theme(panel.grid.minor.x = element_blank()) +
  labs(x="Concentration (ng/mL)", y="Dose change",
       color="Resulting concentration (ng/mL)",
       title="Computer dose adaptations",
       #subtitle="Physician dose adaptations shown as smooth
",
       shape="Absolute dose size (mg)") +
  theme(legend.position="none")
ggsave("computer_doseChanges.png", width=6, height=6)

last_plot() %+% db2Doses +
  labs(title="Physician dose adaptations")
ggsave("physician_doseChanges.png", width=6, height=6)


# interactive dose adapta ----------------------------------
--------------
library(gganimate); library(png); library(RCurl); library(gr
id); library(jpeg)
img1 = readPNG(getURLContent('http://cdn2.iconfinder.com/dat
a/icons/animals/48/Turtle.png'))
img2 = readPNG(getURLContent('http://cdn2.iconfinder.com/dat
```

```
a/icons/animals/48/Elephant.png'))
img3 = readPNG(getURLContent('http://cdn2.iconfinder.com/dat
a/icons/animals/48/Hippopotamus.png'))
imgADAPT = jpeg::readJPEG('ADAPT.jpg')
imgFIT = jpeg::readJPEG('FIT.jpg')
imgMEASURE = jpeg::readJPEG('MEASURE.jpg')

myRaster <- function(raster, data=ggplot2:::dummy_data(),
                     xmin=-Inf,
                     xmax=Inf,
                     ymin=-Inf,
                     ymax=Inf,
                     interpolate=TRUE) {
  raster <- grDevices::as.raster(raster)

  layer(data = data, mapping = NULL, stat = StatIdentity,
        position = PositionIdentity, geom = GeomRasterAnn, i
nherit.aes = FALSE,
        params = list(raster = raster, xmin = xmin, xmax = x
max,
                      ymin = ymin, ymax = ymax, interpolate
= interpolate))
}
ggplot(mtcars %>% mutate(FOO = cut(mpg, breaks=4)), aes(mpg,
wt)) +
  myRaster(img1, data=. %>% filter(as.numeric(FOO)%%2 == 0),
xmin=10, xmax=15, ymin=4, ymax=5) +
  geom_point(aes(group=1)) +
  transition_states(FOO) +
  enter_fade() +
  exit_shrink() +
  ease_aes('sine-in-out') +
  ggtitle("State {closest_state}")

foo <- scenarioA %>% filter(ID==1)
foo$target[1] <- list(numeric())
foo$next_observed[sapply(foo$next_observed, is.null)] <- lis
t(tibble())
#### For the animation, there are 3 states
#### Step 1 (FITTED): What the computer thinks will happen i
n the future
```

```
#### ADAPT THE DOSE
#### Step 2 (ADAPTED): What the computer thinks will happen
in the future
#### ACTUALLY MEASURE
#### State 3 (REAL): What actually happened, while still sho
wing the previous fit
#### FIT THE OBSERVATION
z1conc <- foo %>% mutate(OBS=row_number()) %>%
  uncount(3) %>%
  mutate(STATE = factor(row_number() %% 3,
                        levels=c(1,2,0),
                        labels=c("FITTED", "ADAPTED", "REAL"
))
  ) %>%
  mutate(frame = row_number()) %>%
  filter(frame >= 4) %>%
  filter(frame < max(frame)-1) %>% ## remove last two frame
(REAL)
  ggplot(aes(x=TIME)) +
  theme_minimal() +
  geom_hline(yintercept=13.5, linetype=2)+
  annotate("rect", fill="green", xmin=-Inf, xmax=Inf, ymin=1
2, ymax=15, alpha=0.1) +
  geom_rug(data=. %>% unnest(regimen), sides="b", color="red
") +
  geom_point(data=. %>% unnest(observed) %>% filter(TIME < n
ow), aes(y=Cwb)) +
  geom_vline(data=. %>% unnest(next_observed) %>% filter(TIM
E > now), aes(xintercept=TIME), alpha=0.2) + #show all futur
e observation times
  geom_point(data=. %>% filter(STATE=="REAL") %>% unnest(nex
t_observed), aes(y=Cwb)) +
  geom_line(data=. %>% group_by(STATE, frame) %>% group_modi
fy(function(x, keys){
    fit <- x$fit[[1]]
    regimen <- if(keys$STATE=="FITTED"){ x$regimen[[1]]} els
e {x$next_regimen[[1]]}
    predict(fit, newdata=seq(0, 15*24), regimen=regimen)
  }), aes(y=Cwb), alpha=0.3, color="blue") +
  geom_line(data=. %>% group_by(STATE, frame) %>% group_modi
fy(function(x, keys){
```

```
    fit <- x$iterationFit[[1]]
    regimen <- if(keys$STATE=="FITTED"){ x$regimen[[1]]} els
e {x$next_regimen[[1]]}
    predict(fit, newdata=seq(0, 15*24), regimen=regimen)
  }), aes(y=Cwb)) +
  geom_vline(aes(xintercept=now)) +
  geom_point(data=. %>% rowwise() %>% mutate(real_obs=list(f
it$observed)) %>% unnest(real_obs), aes(y=Cwb),
             shape=3, alpha=1, color="blue") +
  scale_x_continuous(breaks=seq(0, 14*24, by=24), labels=seq
(0, 14)) +
  coord_cartesian(xlim=c(0, 15*24)) +
  theme(panel.grid.major=element_blank(),
        panel.grid.minor=element_blank()) +
  labs(x="Time post transplant (days)", y="Concentration (ng
/mL)") +
  myRaster(imgFIT, data=. %>% filter(STATE=="FITTED"), xmin=
0, xmax=50, ymin=30, ymax=40) +
  myRaster(imgADAPT, data=. %>% filter(STATE=="ADAPTED"), xm
in=0, xmax=50, ymin=30, ymax=40) +
  myRaster(imgMEASURE, data=. %>% filter(STATE=="REAL"), xmi
n=0, xmax=50, ymin=30, ymax=40) +
  geom_text(aes(x=0, y=40, label=factor(STATE, labels=c("FIT
TED", "ADAPTED", "MEASURED")), group=0, hjust=0, vjust=0))
#z1conc
lapply(1:10, function(i) {
  zTmp <- z1conc + ggforce::facet_wrap_paginate(~frame, nrow
=1, ncol=1, page=i) + theme(strip.text=element_blank())
  ggsave(paste0("scenarioA-subject-conc-frame",i,".png"), pl
ot=zTmp, width=16, height=9)
})
z2conc <- z1conc +
  transition_states(frame, transition_length=1, state_length
=3) +
  enter_fade() +
  exit_shrink() +
  ease_aes('sine-in-out')
# z2conc
z2conc_gif <- animate(z2conc, fps=30, duration=60, renderer
= gifski_renderer(loop=FALSE))
anim_save(paste0("scenarioA-subject", foo$ID[1],"-conc.gif")
```

```
)


z1regimen <- foo %>% mutate(OBS=row_number()) %>%
  uncount(3) %>%
  mutate(STATE = factor(row_number() %% 3,
                        levels=c(1,2,0),
                        labels=c("FITTED", "ADAPTED", "REAL"
))
  ) %>%
  mutate(frame = row_number()) %>%
  filter(frame >= 4) %>%
  filter(frame < max(frame)-1) %>% ## remove last two frames
(ADAPT, REAL)
  ggplot(aes(x=TIME)) +
  theme_minimal() +
  geom_col(data=. %>% group_by(STATE, frame) %>% group_modif
y(function(x, keys){
    regimen <- if(keys$STATE=="FITTED"){ x$regimen[[1]]} els
e {x$next_regimen[[1]]}
    regimen
  }), aes(y=AMT)) +
  geom_point(data=. %>% rowwise() %>% mutate(regimen=list(fi
t$regimen)) %>% unnest(regimen),
             aes(y=AMT), shape=3, alpha=0.5, color="blue") +
  geom_vline(aes(xintercept=now)) +
  scale_x_continuous(breaks=seq(0, 15*24, by=48)) +
  coord_cartesian(xlim=c(0, 15*24)) +
  labs(x="Time (h)", y="Dose (mg)")

z1regimen + facet_wrap(~frame)
z2regimen <- z1regimen +
  transition_states(frame, transition_length=1, state_length
=3) +
  enter_fade() +
  exit_shrink() +
  ease_aes('sine-in-out')
z2regimen_gif <- animate(z2regimen, fps=30, duration=60, ren
derer = gifski_renderer(loop=FALSE))
anim_save(paste0("scenarioA-subject", foo$ID[1],"-regimen.gi
f"))
```

```
rm(foo)
```