hello Im Rob bees and Im faculty at

Buffalo and uh I will be presenting part

two of a vignette on population

pharmacokinetic

modeling so in in part one of this
vignette we covered elements related to
the research question the nature of an
experiment the data that were generated
and then mainly the visualization of the

data in part

two were going to now is move into the
actual model building what are the
elements of model contexts to explore
what models that we want to
consider what do we look at in terms of
output and what are some very basic

evaluation and how does how does that allow us to generate inferences about the pharmacokinetic characteristics of

elements of model

the

drug again were going to revisit our kopine example from the N imh funded uh

KY studies

so our general outline will focus on

model building approaches and this
encompasses the criteria for model
discernment things such as objective
function value or of information
criteria graphical
indicators then we consider what a
structural b or base model search looks
like how do we approach generating

initial estimates

what is an appropriate scope of
potential models to explore for example
one or two or three compartment models
linear nonlinear you know the target
mediated drug disposition based
elimination how should we think about
the between subject variability that
arises that we like that that is
critical in a population pharmacokinetic
analysis we wont cover in detail any
evaluation of the residual error
structure in this part

help

patterns in those residuals that would

but one would evaluate any systematic

guide model selection were also going to assess the

impact of uncertainty in parameter
estimates and one may also evaluate
mixture distribution for parameters but
mixture distribution is a little more
advanced and they are outside the scope
of this

vignette we will touch on B basics of covariant modeling so once you establish this base

model you search for correlation between normalized distributions of the individual parameters we call those the AAS and individual specific characteristics uh one could consider continuous versus categorical variables and structures that describe those relationships and then what are the criteria for inclusion or exclusion for the purposes of this part two vignette we will cover the basics of a step high CO variant modeling strategy so lets start with the criteria for model

discernment the objective function so many statistical analyses will return a numerical index of model

Fitness this is often referred to as an objective function value and if youre doing a nonlinear mixed effects modeling population pharmacokinetic analysis for example if youre using nonm that objective function value is based on the likelihood actually the value returned by Nom is minus times the log likelihood and why minus two times the log likelihood well if youre comparing models the difference in minus two log likelihood values is approximately Kai

Square

distributed so what does that mean if
you have a population pharmacokinetic
model you can compare models if theyre
nested with different degrees of freedom
you typically make one change at a time
and you can evaluate whether or not that
change appears to have improved the
model description of the data
statistically by looking at the change
in that minus two log likelihood value
for example if your objective function

its minus two log likelihood if its

points lower with one change in the

model IE one additional parameter then

that means that you have a significant

effect at the P equals 0

level if the objective function value

change is points with one change in

the model so one additional parameter

then its significant at the 00

level so it gives us some guidance its

not the only thing that should be

considered there are many elements that

one must take into consideration when

building

models another criteria for model
discernment is are the information
criteria I have two examples here the
AKA key information criteria which many
people may be familiar with this is a
function of that minus two log
likelihood or that objective function
value but theres a penalty added to
this and that penalty
is two times the number of parameters
that are in the models so if you have a

more complicated model if you add
parameters your information criteria
value will not be as low or as
optimal similarly the basian information
criteria sometimes called the Schwarz

basing

criteria is a function of the minus two
log likelihood plus a penalty that
comprises the the product of the number
of parameters multiplied by the natural
log of the number of data points so here
you have a combination of number of
parameters and number of data
points uh these information criteria do

not give a strict

basis for statistically statistical
inference based on statistical
significance what they do provide is
whether the complexity of the model is
supported by the observed data so its a

binary

decision what is better or worse but not
a specific continuous value that tells
you something is met a particular
threshold with respect to statistical
significance

so continuing along with criteria for model discernment some other basic graphical graphical indicators of

Fitness

include various goodness of fit plots such as the dependent variable versus the population predicted value the dependent variable dependent variable in this case is the concentration measurement versus the individual model predicted value so its adjusted for the individual observations and individual varant and then plots by individual that show the time course of the data show the time course predicted by the model at the population level and show the time course of the predictions from the model at the individual level and then of course residual plots but as I mentioned in the introduction I wont be covering residual plots in uh this vignette uh there are also Visual and numerical predictive checks and again those will not be covered in this particular

vignette so lets go back to the basic
goodness of fit plots for model
discernment the dependent variable
versus PR plot so again the DV dependent
variable are our

observations threed denotes our

population

predictions so theyre B those are
predictions based on the dose time and
adjusted for covariant values if those
are in the model okay and these are
based on the fixed effects that are
estimated in the pharmacokinetic model
the PK parameters and the estimated
coari effects so this is also sometimes

referred to as the

thetas all right Al thetas are not
always strictly associated with fixed
effects they are typically
so so to Aid visualization in producing
this type of plot you want to use the
same range on the X and Y AIS
okay so here we have the observed
concentration and the population
predicted concentration so population
predicted concentration on the xaxis

and The observed concentration on the y

axis the dots are

the uh values for each of the pairs of observe observations with

predictions this black solid line is the

unity

line and this dashed line is a smooth through the data to see whether or not there appears to be a major departure

from that line of

unity its useful for diagnosing

structural model Miss

specification and what youre really

looking for is whether or not theres an

even scatter of points on both sides of

the

line the distances are less of a concern here because those are often adjusted for with other factors when considered

the individual

predictions

okay uh and the again regression line or smooth gives you should hopefully follow

that line of

identity so considering this dependent variable versus population predicted

heres an example that shows a good
model on the left and this is from the
isop model evaluation white paper montre
at all that wases published in CPT
PSP here again on the xaxis are the
population predictions and in the y AIS
are the observations and you see in the

left

panel that we match this line of unity

quite well we have an even

scatter on the right hand panel you see
a missp specified structure and you see
some departure in that smooth from the
line of unity so something something

systematically not being

adequately

predicted or alter L predicted with some

degree of

bias another basic goodness of fit plot
is the dependent variable versus
individual predicted plot again the
dependent variable are the
observations the IAD are the individual
predictions so there predictions based
on the independent variables fixed
effects thetas and the adjustment for

the between subject random effects IE

Adas and

omegas

so conditional on a patients data and

the population

model the model can make an adjustment to try to predict these values more closely and these points should be closer to the unity line as a result uh again to Aid visualization we want to use the same range on the X and Y axis we want to include a line of identity that would be the perfect fit

the right

thats this black line on the plot on

here um and you want to the locally
weighted regression line you can see
that follows for the most part it
follows the identity line and thats
what we we are uh looking looking for uh
specifically again on the plot you have
the observed concentrations on the y
axis and the individual predicted
concentrations on the
xaxis so here we have an example of the
IED versus dependent variable plot for

the true model and the IED versus

dependent variable plot for a

misspecified model and again this comes

from the isop model evaluation paper

Mont at all CPT

PSP and we you see that on the left with
the true model you have a line of
identity the locally weighted regression
line follows that closely the points are
very tightly distributed around that

line of

identity for the true model in the right hand panel with misspecified structural model you can see that points just are systematically shifting around that line

of

identity and that the locally weighted regression also departs from that line of identity so it suggests that you need to rethink your model structure because

there there is some sort of a
misspecification in this model and even
with the adjustments that fundamental
structural model is unable to describe
the pattern of the data Visa the
concentration versus time profile

is this population predicted and individual predicted and dependent variable versus Time by patient or by ID by individual and this is a simple way to visualize how well the model describes individual profiles and its useful for identifying observations that Warren doublechecking for potential mistakes could be about extreme values and that sort of thing so what does this plot look

matrix of individuals so each of these panels represents a different individual we have concentration on the Y AIS we have time on the x axis and the dots are the actual observed data The observed concentrations in this case and we we have two different lines we have a blue line and a red line the blue line is the population predicted value and if we dont have covariates in the model and were all and all the patients got the same dose this blue

line is going to be the same for all individuals okay because its using the pop because it is using the population pharmacokinetic

parameters the red line is the individual predicted lines these are adjusted for the observations so where we can see this most dramatically is an individual 00 you see that they have this one high concentration so the population prediction is this this blue line and it shows the same as every other Blue Line in all these panels but because this solution is conditioned on that on individual 00s data for this particular panel the individual prediction is actually closer you can see that its matching those points a little more uh a little closer at least its trying to match that that that

higher initial

value this allows you to to to evaluate
whether you think there are there there
are outliers whether individuals seem to
be grouped into different kind of
subpopulations in terms of the

concentration time profiles and to evaluate whether or not at the population level beyond the DV versus PR plot the predictions seem to be uh reflecting the data with respect with respect to evenly Under and Over predicting these uh these observed

data

so the next component were to talk were going to talk

about are the initial

estimates and these are basic principles with respect to initial estimates and some of this was taken from a paper in

TCP and I have the citation on on

subsequent

slides so all nonlinear regressions require initial estimates as a starting point for the search of likelihood or search for an objective function and so

this

uh it doesnt matter what software youre using you need to provide inter initial

estimates so theres some strategies for uh considering the calculation or

determining those initial estimates you could do a noncompartmental analysis on the data that you have for example the smooth we showed forthine maybe theres other literature maybe this is a drug that has already come out and there maybe in populations there are uh population pharmacokinetic parameters or even General pharmacokinetic parameters published the FDA uh website in particular the redacted andd uh reviews of clinical pharmacology review in particular are available as a result of the Freedom of Information Act and these these often provide average profiles from phase one and two studies with PK characteristics alternatively the drug labeling the clinical pharmacology section often has an indication of the general pharmacokinetic characteristics and those are all places that you can go to evaluate this now if this is a a brand new drug and youve got your then you you youll have to use your your your the data that are available to you from the study key challenges include

sparse sampling so you know you could do time after dose and consolidate by dosages similar to the plot that was shown in part one for the copine and we could have multiple points create a smooth and do a noncompartmental analysis on that smooth curve at least give us some sense of what reasonable clearance and volume distribution parameters might be so this is often a nontrivial issue as poorly chosen estimates can increase the runtime for any nonlinear aggression problem or any nonlinear mixed effects problem and in particular if youre using nonm and youre using some of the gradient methods and uh the first order first order conditional first order conditional with interaction in theasian it takes longer for solution to be discovered it can also result in poor or incorrect final parameter estimates some of these gradient base methods are quite susceptible to local Minima and if you dont have good initial estimates you may end up with a solution that you

think is a is a stable solution but it actually represents a local minimum in terms of the the the estimated pharmacokinetic parameters so what parameters require initial estimates well our fixed effects values typically thetas right our typical values so these are values in the Nom Control stream and you also probably want to provide some bounds so for example if youre estimating a clearance value clearance is a physiologic process has has to have a positive value uh its probably not zero so you want to put some bounds in and its probably not you know 00 million liters per hour so you may want to give not only an initial estimate but some ranges one would one also must provide initial estimates for the random effects value so for the between subject variability and between occasion variability that you might pose in a population pharmacokinetic model these initial estimates are provided in the

dollar Omega block and it provides the

initial search for a value representing
how variable that parameter is between
individuals or between occasions
depending on which one it is bsv or BV

in the

population and then the residual
unexplained variability and for the
residual unexplained variability you may
want to start with uh values that are
close to the assay variability the limit
of quantitation and the CV of that
particular

assay

So speaking a little more specifically
about fixed effects Theta or typical
values or population averages this is
strictly the central Tendencies so the
population typical values youre not
providing any information on the
variability or dispersion or uncertainty
across the population and again were do
using a best guess and we talked about
strategies for that best guess again
literature is a good source of
information the authors from the TCP
paper suggested some general ranges for

linear processes or typical first order
processes in pharmacokinetics for
clearance and volume and absorption
processes but it really will depend on

your specific

drug uh you know keep in mind that the
values might be higher for the
extravascular dose given the divisor of
f these are really if it is if it is an
extravascular dose its clearance over F
and volume over F

Etc so what about for our

example of

copine well we actually were multiple
values available from the literature and
that this table shows the parameter
values so these were typically for one

compartment

models uh and there were three

Publications in the peerreviewed

literature uh isbister which was

actually in a toxicology context with

overdose Kimco which was an early

evaluation that used the phase one and

early phase two studies to generate the

model and Z was from from actually the

the uh company developed copine with
with with a larger data set we even had
some values from the ktie studies so
these appeared in a chapter reflecting
this of a book that was published from
the studies but were incomplete you see
theres no residual error

provided uh and uh was other limited information so we we have quite a rich

set of information to start
with and we can use those as initial
estimates and also give us a sense of
what sort of model structures we may

want to

explore so what is the scope of a
structural model search well we might
consider the number of
compartments this could be one or two or
three compartments linear mammary models
with first order processes perhaps the
type of an absorption model is this A
first order absorption is this a zero
order absorption is this an llang
distribution with the catary chain uh in
terms of the uh delay in absorption is

there is there a lag in the

absorption in terms of elimination

Pathways is this does this appear to be
a linear elimination or a nonlinear
elimination uh there may be other
contributors to the drug disposition to
the pharmacokinetic profile is this a
drug that is subject to Target mediated
drug disposition what are The Binding
protein binding characteristics are
there Transporters that act on this drug
so lets go back to the profile that we

observed

for the copine study so here again we have a dose normalized concentration on the y axis and the time since last dose on the xaxis and if we look at this remember that there the these appear to be linear processes there theres these are multiple Doses and there are not doses clustering above or below from

vignette um and the general input and
offset look close to sort of straight
lines theres not a lot of not not a lot
of indication of a nonlinear process
here so we probably will explore only

part one of the

linear processes and maybe one and two compartment

models so we thought about okay thats the scope maybe for our basic structural model what about the between subjects variability strategy stry so we have parameters in this pharmacokinetic model and they may vary at random between individuals right because is are between subject variability and there are generally two schools of thought on the approach to this one is to put a between subject variability term on every fixed effects parameter so theres between subject variability on absorption theres a between subject variability onment distribution and theres a between subject variability on the clearance the other school of thought is to put no between subject variability and start with what is effectively a naive pulled analysis or naive pulled minimization and if you go back to the original lecture we discussed naive pooled analyses as one as one potential strategy and also highlight some of the

limitations

there in terms of the structural and
base model strategy if we add the
between subject variability to all fixed
effects

simultaneously we may be able to observe where the bsv estimate is very very small we can decide maybe we should remove it maybe you know a half a percent between individuals is not really going to have any any impact uh now this might also be very small because this the data are so variable or theyre multimodal and the penalty for having an enormous between subject variability estimate is greater than the penalty for not predicting those concentrations as well which is the balance thats thats calculated in the likelihood for these nonlinear mixed effects popul phac kinetic models similarly if the bsv is estimated but extremely large maybe its not a not a single distribution maybe there are multiple modes for example if you have a drug that is eliminated by cytochrome

that have almost no metabolic activity
we know there are intermediate and
extensive metabolizers which have sort
of have the typical elimination
characteristic that are noted for those
drugs and and we also have CL
individuals who are Ultra rapid
metabolizers and theyll have extremely
low level sat concentrations those Z do
not arise from the same distribution
theyre centered in
different positions in terms of the

considerations of off diagonal elements
what are the expected correlations
amongst parameters and we should explore
those blocks structures exploration of
off diagonal elements would be beyond
the scope of the vignette
today so the other strategy is
to not put between subject variability
on in the initial analysis basically do
a naive pulled assessment and then add

bsv parameters one at a time and then decide whether it should be kept in the model do we use the statistical criteria do we do we need to have between subject variability on multiple parameters perhaps theres a dependence that we might miss if we do this again we need to consider nonnormal distributions in multi modal distributions and off diagonal elements so there are various also residuals that we want to consider and you can evaluate graphically using residual uh patterns and Im will to touch on this in a very superficial way we have the Rees residual which is the population residual the W res which is the weighted population residual this is only really appropriate for the first order estimation method the conditional weighted residual which is appropriate for the first order conditional estimation the conditional weighted residual with interaction which

is appropriate for the first order

conditional estimation with the interaction option the individual residual which is the newer versions of NM is now being automatically calculated in the older versions one had to incorp break this into the error block and the individual weighted residual and the waiting depends on the error structure that you select helpful visualizations include the population residual versus dependent variable this will tell you right away if you cannot use an additive model youll see a fan structure this with increasing uh distance from that unit from the zero line with increasing dependent variable the conditional weighted residual versus dependent variable conditional weighted residuals versus time individual residuals versus dependent variable and the individual rated residuals versus the dependent variable so lets look at some model output so we talked about selecting model structures and having a scope of search so this is a game this is from our analysis of the copine data set from

the KD studies on the left table we have output from a one compartment model analysis linear absorption linear absor for absorption in first order elimination on the right panel is a two compartment model for the same data set the numbers are the parameter values you see the units are provided here for the thetas these are the fixed effects okay these are your typical values the population value of clearance 0 for the one compartment model for the two compartment model for example we have volume absorption rate for the two compartment model V V because its an extra value dose and this is the coding in in Nom and then the interc compartmental clearance is Q the omegas are the between subject variability on those parameters so you can see that the we have the Omega clearance right this is the between subject variability on clearance the value returned by nominum is the actual when we convert that to a percentage between subject variabilities

about so the values in Brackets are the

percentages between individuals

between individual variabilities

again these are the between

individual variabilities and in Brackets

we see the actual

percentages and the numbers returned by n are presented as the main numbers here

these are basically like a CV not
exactly depending on which estimation
method youre using you have to do a
different back transformation
and then we have our residual
unexplained variability this is our s

these are these Sigma values and at the very bottom we have our ofv the

objective function value and you can see for the one compartment model this is

the value of

0 9 and you remember that the objective function value is minus two

times the log

likelihood and the lower the better
these values if we compare this to the
two compartment model

we see that this objective function value is one is

00

distribution B and the
intercompartmental clearance or you
could argue four degrees of freedom if
you are including the between extra
between subject variabilities but the
0 points on a Ki Square even for four
degrees of freedom is highly significant

adding these uh extra volume of

so

this would point us towards selecting a
two compartment model structure now this
is in conjunction with evaluating
graphical patterns uh residuals
Etc now were considering between
subject variability you see that the
estimate for the between subject

variability on the peripheral volume
distribution here Omega V is extremely
low right 0 between individual well
that means its probably not having a
significant impact on the disposition

and its an extra

parameter so if we basically remove this

parameter fixing it to zero in the right

side side table here we see that the

objective function value actually got

better so the estimation was more

stable it was and the other parameters

actually are all reason reasonable as

well and youll see this V value

actually starts to get a little bit

lower and closer to what was reported in

the

lature well what about model uncertainty

we can also

consider the parameters with respect to
the model uncertainty this is often also
referred to as the relative standard
error of the parameters its basically
how well do we know those parameter
values we do an analysis and we
determine okay

well can we rely on that parameter
estimate so youll see that most of
these value these are these are the
values in Brackets or parentheses next
to the main numbers here in both of
these tables

and what we see is that theyre
generally relatively low with a couple
of exceptions but the major exceptions
we left in now this is the between
subject variability of Omega V okay we
have a value of in this case but
the uncertainty in this estimate is
000 plus which suggests that that
that value could assume just about any
any number any value and you would get a

very similar solution so these

particular uncertainties or relative

standard errors are calculated using

something called the fisser information

Matrix what is the fer information

Matrix if youre running most nonlinear

software a variance covariance Matrix of the parameters is determined that Matrix is inverted and it allows one to make an

mixed effect modeling

inference about how much worse the model fit gets if you change the parameter value and that is the basis of the determination of these uncertainties in these parameter values in very rough terms in very rough terms uh one can also obtain these values using a nonparametric bootstrapping approach but thats outside the scope of of this vignette so essentially if we remove again this between subject variability we really get no change in the objective function and the model stays about about the same and estimates are all reasonable so theres no reason to include this so the uncertainty in this case on this parameter estimate pointed Us in the direction modifying the model OKAY by removing this this element so lets talk a little bit about basic covariant modeling and the technique Im going to touch upon in this part two vignette is stepwise covariant model building and this

involves steps called forward addition

and backward elimination so you start with our base model we talked about building the base model and and making those selections and then you test the covariant to determine which one improves the model most significantly so you test each covariate individually you get a set of objective function value changes or differences and you evaluate which ones are significant and then rank order which you know the level of significance how many points is this change by and you can add this covariant to the to the model starting with that most significant covariant and you add the second most

covariant and you add the second most

did it get did the model get better or

not if it didnt get better then you

remove that that second goarant you had

the third one you check did it get

better or not oh it got better so you

leave that one in for example and you

keep doing that until you have no

further

Improvement then you have this full model with coats and then you start

removing

covariants but in a different order so you dont start with the you dont remove the last covariant that you added you add you remove a covariant maybe the most significant covariant that was added during the forward Edition and you evaluate whether or not the model Fitness got worse IE that objective function value went up and if go and typically you have maybe set an alpha threshold for forward addition of 00 so that would be points and in the backward elimination phase you change that threshold maybe to 00 go points you keep on doing this until no further worsening is observed and you have your final model now this is sort of the most basic technique there are lots of other techniques but again they they are outside the scope of uh of this second

techniques but again they they are
outside the scope of uh of this second
vignette what are the strengths of
stepwise cant modeling well allows
selection of cant relations from a large
set of candidate relationships it has a

reasonable predictive performance theres lot of experience with this approach iss an economy in which models to test and avoids many problems associated with search based on empirical base estimates for example it allows to you to account for varying amounts of information between subjects time varing covariant and a single framework for model selection like oh this is terrific except there are some substantial weaknesses here theres no guarantee that youll find the best model theres a selection bias theres a an increased risk of false selections for relations with lower power and if youre covariant right with power to detect the coar less than 00 And The coefficients are generally overestimated IE the impact on the parameter value and the confidence intervals are underestimated youre doing all these different tests and youre getting different sort of differences and objective function value and Crossing

thresholds but often multiple testing penalties are not applied and we do dont know what the appropriate penalty would be to apply in this case uh its not designed really for generating predictive models and many runs may result in unsuccessful terminations or local Minima so stepwise Cate modeling is shown in this figure here starting with a base model looking at diagnostic information test all the suspected covariates by adding them to the base model independently one at a time pick the model with the most sign ific Cate add the CATE to the base model okay then you repeat come back and say take take the second most significant and add it did the model can get get even better if it didnt remove if it did keep this and keep cycling until you have your forward Edition model there at the bottom of the

First Column of boxes

here and then our backward elimination
you test the effect of removing all
covariates independently one at a time

maybe start with the most significant covariant from the model just dont do the exact same order that or verse of the same order that you did in the

forward

direction and you end up eventually with the new coar based model and you keep repeating this removal until you get no further worsening in the model and you

have your final

model so to

summarize uh our model building vignette

we have multiple

considerations we have to come up with

rules for Discerning what we think a

good model is uh we have numeric indices

to help guide this things like objective

function value Val information criteria

parameter uncertainty there are

graphical indicators such as residuals

and uh these various plots we

discussed uh we have to come up with

initial estimates we can go to the

literature or do a noncompartmental

analysis of average profile over the

population from data that been

generated uh we have decide on a scope of the model search what is an appropriate space to search in terms of number of compartments and different processes linear nonlinear elimination absorption processes how do we handle the stochastic elements of the model the between subject variability do we put this on every parameter to start with or do you one at a time there was a dual unexplained variability what is an appropriate model there is it additive is it proportional is it a combination of the two and then we start to consider covariants how do we incorporate these how do we look at potential relationships and what does that mean are we in danger of missing certain uh uh dependencies amongst covariates and as we you saw from the part one exploratory data analysis vignette one has to be very careful about covariates that are correlated amongst different arms or uh even within different characteristics because what is the characteristic that is driving the

relationship with the response and if you have correlation thats very very difficult to tease apart so there were some elements that were not covered in this part two vignette that is a more specific evaluation graphical evaluation of residuals bootstrapping technique visual predictive checks numer numerical predictive checks and numerical predictive distribution errors and that would be the subject of future vignettes again I would like to acknowledge Kristen BOS at Johns Hopkins University a faculty member there in the in the division of clinical pharmacology and uh an MS candidate in the lab haranga who has been uh working on the the M the model Direct me thank you very

much