our next speaker is dr robert bees who

pharmaceutical science at state

university of new york at buffalo

prior to this dr bees was an associate

professor at indiana university school

and director of d of disease and

therapeutic response modeling program at

the indiana clinical translational

science institute

dr bees received his bachelor of science

degree in pharmacy from the university

of toronto

his doctorate of pharmacy degree from

the university of texas and his phd in

pharmacology from georgetown university

this was followed by a postdoctoral

training at the center for drug

development science until 000

dr bs research focus on the application

of pharmacometric approaches

please enjoy todays lecture

so im going to take the next

about 0 minutes

to introduce some principles of

population pharmacokinetics

i would like to thank the organizers for

the opportunity to

present

these materials to you

so population pharmacokinetics

what does that

mean to uh to uh to to to people

well it has a long history

actually much of it originating

here at

the national institutes of health

and

was really pioneered

by

louis shiner

in his in his

desire

to optimize patient therapy

you can see that

there are

papers beginning in the late

90s

relating

to the use of

computational aids

for

precision medicine effectively

and the initial efforts were related to

anticoagulation therapy

this work

uh continued

to evolve as

uh dr scheiner moved to the university

of california san francisco

and continue with a continued interest

and

computeraided drug dosage

often also referred to as therapeutic

drug monitoring

an important foundation is this

computational component is the

development of population models

these initial

efforts though

did not

incorporate key

components of population modeling and

the understanding of repeated measures

and within individual correlation

that would be essential

to the development of population

pharmacokinetic models

in fact

a seminar by lou scheiner at ucsf

this is a story that was relayed by by

carl pack i believe lou shiners first

fellow at ucsf

that he was introducing

his

work on optimization of therapy

and had bar rosenberg whos an

econometrician come from uc berkeley to

hear this talk

and introduce him

to techniques used in econometrics

that he thought would be

readily applicable

to

repeated measures of drug concentrations

in patients

this work continued to evolve you can

see that theres a computerassisted

digoxin therapy paper that came out the

new england journal medicine in 9

with carl peck and lou shiner

in addition

uh this work continued to to to involve

to to to evolve

incorporating individual specific

measurements in this 9 paper that now

you see includes bar rosenberg

the first

real estimation of the population

characteristics though in what would

become the software nonmem

uh was published in 9 and this is the

estimation of population characteristics

of pharmacokinetic parameters from

routine clinical data

so thats some some background and you

can see that a lot of the thinking in

the foundations are not particularly new

but they still have

i think major applicability

so i guess

when i say a population model or a

population pharmacokinetic model

what are the key features

that

were trying to get at with a population

analysis or a population model in

general

so the first

is understanding the central tendency

of a population

what describes

a typical

atypical response

a typical concentration time profile for

example

a typical elimination rate of a drug

some basic descriptors of central

tendency include the mean the arithmetic

average of all the available values of

the descript of a descriptor where the

descriptor is a random effect

the median the descriptive value such

that 0 of the remaining values are

larger and 0 are smaller

and the mode

the descriptor value at which the

probability density attains its maximum

value

ie highest point in the distribution if

youre looking at a maximum likelihood

its the highest point of that

distribution

these are not the only measures of

central tendency

but i think are some key ones just to

set up

the thinking with respect to

what appropriately describes

our group

the other component

of a population model

is understanding the variability in that

group

so its one thing to know on quote

unquote on average

what is a response

its

quite

quite an extension of that and i think a

critical extension to understand well

how variable is that response

and how does that variability distribute

across

individuals in a group

and so these variabilities are also

described by a variety of statistics and

its up to the

user its up to the investigator to

decide

what the most appropriate descriptor is

here i have listed a few very basic

examples

a standard deviation square root of the

variance

the variance the average of the square

deviations of the random effects from

the mean

ie the differences between each

individuals response and your measure

of central tendency

in this case

if its a variance this is the

deviations from the mean and squaring

these deviations to exaggerate

to exaggerate deviations as they move

further away from that central tendency

the coefficient of variation which is

simply the ratio between the standard

deviation and the mean

and quantiles so percentile

distributions if we have enough

information to understand the

distribution what are the quantiles

where is the central 0 percent of the

population

how what are the responses expected

between the fifth and 9th percentile of

the population

what are the responses

expected in 9

of individuals for example

so

what are some general approaches to

population modeling

now much of this sequence is taken from

a

9 series of papers by lou shiner and

stuart beale

on

population

approaches

i have

six approaches listed here

and were going to focus or were going

to go through actually the first five

and focus really on the fifth

fifth approach

these general approaches include naive

pooling

fitting the average profile

a standard twostage and iterative

twostage approach

nonlinear mixed effects modeling and

bayesian inference

so

lets begin with naive pooling so what

is naive pooling with naive pooling all

data points are assumed to arise from a

single individual

you fit a single function

to this combination of all individuals

so you ignore the interindividual

variability as well as correlation

within individuals

so you really are lumping what should be

two hierarchies of variability

were getting basically just a residual

a residual distance thats estimated by

whatever choice of estimator youre

using would be these squares or maximum

likelihood or what have you

so

in this slide we have uh

five sample

concentration time profiles

denoted

by these different

colored symbols we have concentration on

the yaxis and time on the xaxis and

these look like typical

first order absorption first order

elimination

pharmacokinetic profiles

all right

if we do a naive pooled fit

we get this solid line you may say well

that looks reasonably good

even though weve ignored the between

subject variability in the context of

not

considering the within individual

correlations

specifically across each of these

simulated

individuals

the next approach is fitting the average

profile

so data points in this case are averaged

at each point

in a measurement sequence

that means that your measurements have

to be made at exactly the same time

across all individuals

it also means that you have to have

measurements for all individuals at

those times you start to have missing

information

this approach

cannot be used

the average points

across time are then used to fit or

produce estimates for the

pharmacokinetic model so returning

to our simulated

example

here again we have concentr the graph

showing concentration on the yaxis and

time on the xaxis

and the blue points are the average of

all of the concentration measurements

across the five simulated individuals

okay

and so what were doing is were fitting

a pharmacokinetic model to this average

so were effectively ignoring the

individual contributions

were ignoring the uncertainty of the

individual contributions were saying

that the average at any given time point

for that concentration is a sufficient

statistic

to describe that

profile

and if we do a least squares or maximum

likelihood fit we get this pink line

which looks like it goes through the

dots reasonably well maybe it

underestimates these last these last few

points i said well that may be

sufficient for me

so why not

use naive pooled or averaged approaches

when we talked about principles of

what comprises a population model

we can only get at the central tendency

with naive pooling or fitting the

average curve

and

as we see from

this table

we have the

true values or the values from which the

pharmacokinetic profiles were simulated

in this case a clearance of 0

a volume of 0 and an absorption rate

constant of 0

and we see that the estimates for naive

pooled and average fitted

are systematically off even though the

curves fit reasonably well

if we were interested in retrieving a

parameter that perhaps had some

physiological relevance

we would get a parameter that may be

skewed

of course this point is argued back and

back and forth but well take that

position

for the sake of todays lecture

so in this case you see the true value

of clearance is 0 and were

underestimating this at about

being

the return result for both naive pooled

and average fitted so weve

underestimated the clearance

about percent of the true value

theres a corresponding increase in the

estimate of the volume of course this

compensates and gives you approximately

the same curve

the value that was used for simulation

was 0

the naive pooled value is 0 and the

average fitted value is approximately

0 as well

similarly for the absorption rate

constant

we simulated

from a true average of 0

and we see that the naive pooled value

is about 0 percent higher as well as

the average fitted value

so if were interested in retrieving

even the central tendencies

there appears to be

the possibility that you could have bias

in that response and weve not addressed

the issue

of variability

or between subject or interindividual

variability in this case

so our correlations within individuals

are ignored

so how might we start getting at

the issues of variability

so theres a standard twostage approach

if you have a complete pharmacokinetic

profile for all of your subjects we do

in our simulated case for five

individuals

and we have relatively intense sampling

which means we have a lot of

concentration measurements for each

individual that allows us

to estimate the pharmacokinetic

parameters for each of those individuals

okay

you then take those individual estimates

calculate a sample mean and variance

that becomes your population central

tendency

population average and your population

variability is that covariance

or variance of the parameters

in this slide we show each of the five

individuals

concentration time profiles separately

these are the actual the dots are the

actual observed data

and the pink lines are the

predictions of the model

that is fit to that using this standard

twostage approach

and you can of course obtain those

estimates using any sort of nonlinear

regression

package

any sort of ordinary least squares or

these squares or maximum likelihood

approach

so these profiles look like theyre

reasonably well captured

just to recapitulate the principles of a

standard twostage approach

the model is

identified separately

in all subjects for example with least

squares so you get the individual

estimates of the parameters okay

and then we move to the right panel the

population mean and all of a sudden im

introducing this theta here

theta is typically used to denote a

population central tendency

in

uh population pharmacokinetics

and the covariance omega squared

are calculated

and you may wonder why im using the

term covariance well the covariance of

this of of of this of a parameter with

itself is the variance so this allows us

to use this as a general term

to understand only

the variance the population variance

but

how parameters might covary okay

but requires richly sampled data

so you have to get those estimates in

each

individual

okay

and

it ignores the precision of the

individual estimates

so you get these values

but you dont really you you take a

central tendency and a variability but

you dont know how well you know that

central tendency and that variability

and there were several papers one by ted

grisella i believe its almost 0 years

ago now showed that actually you you may

overestimate the population variance

perhaps as an artifact

of

not propagating the uncertainty in the

parameter estimates into those summary

statistics that youre calculating

so standard two stages actually closely

related to iterative two stage youll

see that the top part of this panel is

identical

to uh the previous slide effectively so

you identify the mod model parameters

separately in all subjects all subjects

you calculate a mean and variance from

all the individual estimates

and then

you take those mean and variances and

you use them as the new starting values

for the next analysis in each of those

individuals okay it uses an empirical

base prior

an individual estimation and then you

reestimate the parameters

take the mean and the variance again or

covariance

and

you are actually

uh attempting to refine those estimates

and would

com

would repeat this cycle

until your estimates of the population

mean and population variability

are stationary theyre not changing any

in any further

it can work in some sparse data

situations

and it tends to give more reliable

individual estimates

so how did the

standard twostage

approach

do

with respect to retrieving

the parameters that were used for

simulation

for these for this population

pharmacokinetic example

so again

here we have our pk parameters clearance

volume and ka

the true values of 0 0 and you can

see that the standard two stage

retrieved those parameters exactly

okay

similarly we now are getting at

population variability in this case

were saying the standard deviations

parameters reflects

the between subject variability in this

case

all right

more specifically

you can see that the true population

variability and clearance was the

standard two stage returned

the

standard deviation for the volume 09

we got 09 with a standard two stage

and

0 for ka

point you can see that we basically

retrieved exactly the same values

so of course were not showing inflation

here in the between subject variability

in this particular example

so lets think about a distribution of

clearance valves lets get back to the

idea of

variability in a parameter okay

so

clearance values are usually assumed to

be normally distributed or log normally

distributed so lets say normal in the

log domain okay

centered at zero so zero means centered

at whatever the population central

tendency is so you do a normalization so

that your deviations are

uh relative to a specific distribution

at zero

okay or

about zero okay

and i also wanted to introduce some

notation here so that you get more and

more comfortable with this so here this

is effectively a gaussian distribution

this histogram with probability on the

yaxis and normalized parameters

distribution for clearance gain centered

at zero

with normalized values here okay

because its a normal distribution

youll see that our central tendency

measures are equal a mean is equal to

the median is equal to the mode

right

and so that tells us okay thats uh

thats our p thats the piece of our

population model that tells us about

typically what is the population

response likely to be

but we also

are illustrating

the spread

of that probability density okay and

this probability density can be measured

by standard deviation or variance and it

tells us how variable individuals might

be one to another in their particular

response in this case how variable are

the clearances across the population

additional notation youll see the

standard deviation of course the square

root of the variance so were denoting

that omega

and often in population liter in the

population literature between subjects

variability is denoted

with

the term omega omega squared is our

variance

okay

and here we have

another term heres an eta so eta is the

individual

deviation from the central tendency okay

and were going to say for the sake of

argument today

that this eta is drawn from a normal

distribution thats this large n

centered at zero

with variance omega squared

so we know now that we can make

adjustments

for individuals okay we can make it on

the basis of knowing

what our population variability is

and that we have this type of a

distribution so that we can its

centered at zero

okay

this is also of course the gaussian

distribution also known as the bell

curve

all right

so lets now

think about nonlinear mixed effects

modeling that was the next one on our

list

okay

we want to go

weve so far looked really at at

modeling a population with naive pooling

with

fitting the average curve

we fit in separate individuals okay

with the standard twostage and

iterative twostage approach

and then come back to the population but

lets go from the individual to

population and add an additional

statistical layer to this a statistical

model to this okay so lets say we have

many sparsely sampled individuals we no

longer have intensive sampling

that forms a population okay

we want to be able to appreciate the

population response and if we have

sparsely sampled individuals initially

we can only get at the population

response population central tendency and

variability okay

however so we have to introduce the

statistical model then to get at this

and

that allows us to characterize

and model sources of variability so we

can start to explore what may explain

the variability

in a response the variability in the

elimination rate the clearance the

variability in the volume of

distribution the variability in the

absorption rate

okay

when we establish the population model

though we can actually go back to those

individuals even if theyre sparsely

sampled and get some additional

information

that tells us what the most likely

parameters might be for that individual

so we start with many sparsely sampled

individuals

we develop a

population model population response we

understand what that central tendency is

what the likely variability is

and then we can condition on that model

we can use that model to get at what the

most likely

response is for that individual with

respect to their elimination rate

constant or their clearance or their

volume of distribution or their

absorption rate

in addition

because weve started to define

variability

we can begin to explore

explanatory variables that may

help us to understand

deterministic

contributions to variability

from one individual to the next

so this entails

mixed effects okay this is adding an

additional stochastic model

and this stochastic model tries to get

it what is the extent of variability in

the pk parameters between subjects

between subject variability

you might say well didnt we do that

with standard twostage

we took the average and we took the

standard deviation we said the standard

deviation is our description of the

variability yes in a sense we did but

with a standard twostage or irritated

two stage you cant use sparse sampling

and you ignored the uncertainty in that

in those individual estimates in

defining

these these additional hierarchies of

variability okay theres actual levels

of variability

because we can also now get at

what is the extent of the variability in

model parameters

in the same subject studied on multiple

occasions so you have a patient hes

coming back multiple times

how big is that variability at random

from occasion to occasion

and you can imagine we can add multiple

additional hierarchies of variability

depending on what youre trying to

explore with your model

im just going to explore these three

the last one being the residual

unexplained or residual unknown

variability

this gets at

what is the extent of model

misspecification and unexplained

variability in the concentration effect

measurements

okay so even though we can now adjust

for

random between subject variability

and within individual

between occasion variability

we still dont perfectly predict an

observed data point for that individual

or data points for those individuals

concentration measurements for those

individuals

and that has to be described by

a variant structure and thats that

residual variability component

so weve talked about mixed effects and

we

weve talked about hierarchies of

variability but

what are these mixed effects the mixed

effects comprise what are known as fixed

and random effects

now something to keep in mind is is if

youre doing a bayesian analysis a

purely bayesian analysis lets say

youre implementing this in a software

like stan or like bugs or even in

nonmem uh and its a pure bayesian

analysis

all of the elements are random effects

just so that when youre reading the

literature you wont you you youll know

that that you can you have to parse what

those what those variables are for the

sake of

todays lecture im going to focus on

the area of nonlinear mixed effects

modeling where we really separate out

fixed and random effects

and fixed effects are features common to

the entire population

so in this case

we have clearance volume of distribution

and the absorption rate constant

every individual is going to have some

absorption rate constant

theyre going to absorb the drug right

every individual is going to have some

volume into which the drug distribution

distributes

every individual is going to have some

rate of clearance some elimination rate

of that drug

so theyre common to the entire

population okay

now you may have individual specific

characteristics that interact

to help describe systematic

contributions to differences in these

parameters

and these are the covariates these could

be things like

body mass index

could be specific genotypes that are

related to

uh the

that are related

to

the

expression of metabolic enzymes that are

specific for your drug for example or to

transporters related to

absorption efficiency of a particular

compound for example

and then we add the population

variable variability and measurement

uncertainties as our random effects on

top of this so our fixed effects

features common to the entire population

alpha interval interact with covariates

okay

and then we have our random effects this

is our population variability and

measurement uncertainties this is the

between

subject or intersubject variability

between occasion variability or

interoccasion variability

and our residual unexplained variability

and here ive just illustrated that

heres our eta again we have a to i this

time because were sampling for the ith

individual from this distribution and

this distribution this

absolute symbol here is basically saying

arising from or drawn from

a normal distribution so showed the bell

curve previously centered at zero with

variance omega squared

okay

so that that is sampled with a varying

probability depending on how likely that

is in this how likely a parameter is

in this particular population once you

understand the central tendency and the

variability of the population

so we can consider this a model of

physiologic parameters not to be

confused with physiologically based

modeling but a model of physiologic

parameters this for the sake of argument

and its often denoted

as

the random and fixed or varying plus

constant components

of a

of a model

its a combination of an input output

model and a stochastic model so the

input output model is our classic

pharmacokinetic model it could be our

first order absorption first order

elimination

uh

model

and the stochastic model are the

elements that introduce the between

subject variability between occasion

variability residual unexplained

variability

so i think weve weve talked about this

but well perhaps recapitulate types of

variability or stochasticity that

represent that are represented by random

effects

again between subject variability often

referred to as between individual

variability or intersubject variability

between occasion variability or

interoccasion variability so how

variable is this individual from

occasion to occasion in their

concentration time profile that leads to

different or at least estimates of

clearance values that may be different

from occasional occasion and a residual

unknown unexplained variability ruv also

often referred to in the literature as

interest subject

variability intraindividual variability

and or within subject variability

although strictly speaking

its related probably more to modern

specification and and uh error issues in

in

in in in the ascertainment of the

measurement

so lets talk in a little more detail

about between subject variability

so

the between subject variability is our

quote unquote statistical mathematical

expression relating subjectspecific pk

or pd model parameter values to random

variables describing the population

variability

okay thats rather

abstract perhaps

so lets think about it in more concrete

terms so here we have an equation

where this clearance i equals clearance

pop plus a to i so what does this denote

clearance i is that

a clearance for a given individual

okay

and the clearance for that given

individual is a function of the central

tendency for the population the

population clearance

plus an adjustment factor heres our

adai showing up

and that adai is sampled being drawn

from

selected based on

the between subject variability

distribution that we defined previously

centered at zero with variance omega

squared

okay

and that variance omega squared

is going to

constrain

how far

that individuals clearance value can be

from the population value

so if the between subject variability in

the population is very large

then you can pick lots of

highly varied

values

okay to make that adjustment when youre

drawing a to i

okay or when the software algorithm is

deciding on that a to i okay

if youre between subject variability in

the population as

use as you have estimated it

is small

then the improvement in the description

of that individuals data is going to

have to improve quite dramatically in

order to justify

selecting an adjustment factor and ada a

an individual clearance value that is

substantially different from the

population value

okay

so below here ive just indicated again

that the clearance eye is the

individualized value of drug clearance

clearance pop is a population clearance

and that eta is from the random effect

is usually drawn from a normal or

gaussian density with a mean of zero and

a variance omega squared and heres this

notation again eta drawn

from a normal distribution centered at

zero with variance omega squared

okay

so lets look at this sort of a

in more of a visual

visual way okay so here we have

two subpopulations males and females

concentration on the yaxis time on the

xaxis

the blue line is the population central

tendency the population average

for

a male concentration time profile

the red line

is the population central tendency

for the

female

concentration time profile so these are

concentration time profiles arising

from

the population average clearance volume

and absorption rate for males and

females respectively okay

its not the average of these other

curves

these other curves are showing the

individual profiles

okay so youve got

three male

representative male subjects shown here

and three representative female subjects

shown below in the red dashed and dotted

lines

you can see that they distribute across

the central tendency

so you can estimate this estimate the

central tendency but none of these

curves may look exactly like this

population central tendency

okay

and these are different

these are different based on the

interest subject or between subject

variability

in this case i have this notation

indicating here why

why well why why in this equation is the

concentration observation

okay you notice we dont have any

residual variability yet and that y is a

function of time

theta remember theta showed up before in

that slide in the standard twostage and

iterative twostage analysis slide

that is our vector of fixed effects

what is our vector of fixed effects

that is our clearance and our volume and

our ka

for the population in this case

separates for males and for females in

this case

in addition we now have this eta we can

sample from to get individual values for

the absorption rate for the volume

distribution for the clearance and those

are sampled from some

normal distribution centered

zero with variance omega squared and it

could have separate variances for each

of the parameters

okay

similarly for between occasion

variability

again if we go back sort of to the

platonic realm right the abstract realm

its our statistical mathematical

expression

relating subjectspecific pkrpd model

parameter values to random variables

describing their variation between

different study sessions

okay

well lets get back now to the practical

implementation here

what we have here is this equation that

says the clearance for the ith

individual on the kth occasion

okay

is again a function of our population

clearance

you see the population clearance is

adjusted by the same

adjustment factor for interindividual

variability theres only one

right one eye represents the individual

but in addition

every occasion where a patient would

return and get a concentration time

profile measured or several

concentration measurements measured

is a new occasion a cath occasion

occasions one decay

and you have a separate adjustment

factor now

that actually modifies

this value

depending on the occasion and

conditioned on the data that are being

observed at that occasion

all right

again

clearance ik is the value of the drug

clearance for the pk model

for the ith individual on the kth

occasion

the population clearance again is that

population value of drug clearance

theres one value

for the population okay

ada

is our between is our sample of our

between subject variability random

effect again drawn from a normal

distribution centered at zero with

variance omega squared

and

kappa here is our between occasion

variability random effect again were

going to assume that its a normal

distribution it doesnt have to be

centered at zero

with variability pi squared

okay

so

a different variability distribution

and its sampled multiple times

okay every time theres an occasion

within an individual

the between subject variability is

sampled once this is how this individual

this is the most these are the typical

values for the for this individual with

the highest likelihood

and then at each occasion theres an

additional adjustment

for

the data observed on that occasion or

the concentration time profile for

example observed on that occasion

so lets

add this into the model now our

stochastic model has between subject

variability and between occasion

variability

again y in this case is our

concentration observation

okay

and its a function of time

theta so our vector of fixed effects

again thats the population average

value

for those in for for for that population

or that sub those the the those sub

populations for clearance absorption

rate and volume of distribution

and then we have a between subject

variability that is sampled from a

normal distribution centered at zero

with variance omega squared and if we go

to the bottom left panel you can see

this is the same panel shown from

between subject variability slide

central tendency

for female central tendency for males

and individual profiles for males and

for females you can see they vary thats

based on between subject variability

if we take one of these profiles

okay

and we propagate it out for multiple

occasions now we can see the effect of

adding between occasion variability the

bottom right panel again has

concentration on the yaxis and time on

the xaxis

the black line is if there was no

between occasional variability you can

see repeated dosing and youre pretty

much at steady state at the second dose

here

if we have substantial between occasion

variability this is the dashed line the

first curve tracks

and then the second one you can see that

the concentration time profile for the

same dose is quite different and then

for the third dose quite different again

this becomes very very important to

capture explicitly

if you are interested

in potentially targeting a specific

concentration for a drug if you want to

implement therapeutic drug monitoring

if you do not

model the between occasion variability

and get a sense of

the

magnitude of the between occasion

variability

you may discover

that that is very difficult to optimize

if that between occasion variability in

fact is greater for example than your

between subjects variability for

whatever reason

and maybe there are systematic

contributors but maybe these are things

that are not captured

in in the study

if you try to target a concentration

based on this first

profile you get a set of parameters for

that individual a clearance a volume of

distribution and an absorption rate that

would give you

this second

black concentration time profile black

line here

however

given the between occasion variability

system what you would observe is

actually the dashed red line and youd

see that you would actually get a very

different

concentration time profile and similarly

this third profile differed again

so now the next hierarchy of variability

residual unknown variability

this is a again a statistical

mathematical expression related to pkpd

or disease model predictions to the

actual trial measurements or

observations

you remember that the previous profiles

did not show any deviations from those

predicted curves

so theyre really simulated from the

model with between subject

and or between location variability

but no residual variability and we know

that there are deviations between the

observed concentrations that you will

be evaluating if youre doing pkpd

modeling

and

the model predictions and these are

accounted for with the residual

error model so here again y time is the

observed data

is a function of time remember our

function so this includes our vector of

fixed effects or between subject

variability and are between occasion

variability maybe covariate effects

plus

a residual variability term

in time sometimes its correlated with

time sometimes

and all

what we have is

uh the value of the observation again

the value of our model prediction

and the discrepancy

between how well the model does and what

weve observed so what the model

predicts versus what we observe okay

now were going to make the assumption

that this that these deviations

are normally distributed so epsilon is

like that individual sample its the

individual deviation between the model

predicted and the observed value and the

observed value for concentration drawn

from a normal distribution centered at

zero

with variance sigma squared

okay and sigma squared may or may not be

a function of time it may or may not

correlate with other random effects

or interact with other random effects

here are a couple of certainly not

exhaustive list of but a couple of uh

sort of typically used residual error

models

the top one is an additive model

where we have our model prediction with

an additive discrepancy so this would be

a constant difference

so if youre talking about

concentrations this would be actually an

absolute concentration value maybe five

nanograms per ml for example

the second is a proportional error model

structure where we have a

where we have our observation as a

function of time as our related is is a

function of

all of our

pharmacokinetic parameters between

subject and between individu between

occasion variability

and then a

constant proportional error model this

would be like a constant cd model plus

or minus 0 percent plus or minus 0

or you can have a combination model

maybe it actually has an additive

component and a proportional component

you can imagine maybe this is something

similar to a limit of quantitation and

this second term is related more to the

cv once youre the coefficient of

variation once you are in

the

uh measurable range

for those concentrations

so how does this look

when we add it to our

model okay now were adding our residual

unexplained variability our observation

is a function of time

our vector of fixed effects

our kas clearances volumes of

distribution well actually this is this

is

this is the ka clearance and volume

distribution for the population our

adjustment factors for between subject

variability

in this case eta drawn from that normal

distribution

between occasion variability kappa drawn

again from a normal distribution

and now weve got our residual

unexplained variability

okay

this figure is showing concentration on

the yaxis and time on the xaxis

the center curve is our population

average concentration time profile

arising from our population average

parameters

the green and yellow lines are arising

from different sets of individual

parameters adjusted for between between

individual variability between subject

variability

and the green dots are what we actually

observe and that

incorporates this residual unexplained

variability or residual unknown

variability so you can see that it

we get this

these values now that our y values look

like real data points that weve now

accounted for all these hierarchies

of variance

so

now that weve talked about

nonlinear mixed effects models

and

these this this particular approach

uh how did this approach do relative to

the other population modeling

approaches given our our example of five

simulated individuals with relatively

intensive sampling

so again this table is showing the pk

parameters clearance and ka the true

values here of 0 0

again standard twostage retrieve these

values

heres our nonmember analysis or

nonlinear mixed effects analysis

and we got wait a minute we got

thats were were a bit low here about

0 low

0 for volume so weve got the same

volume and weve got a lower ka value

so youre probably thinking well what

could go on what could be going wrong

this is supposed to be the optimal

approach

and yet it did not retrieve the

simulated values

but keep in mind

that the nonlinear mixed effects

modeling

approach uses a maximum likelihood

the descriptor it is using for the

population central tendency is not the

mean of the population

as it is in the standard two stage as it

was from our simulation example

its using the posterior mode

of that likelihood so what the mo what

are those most likely parameter values

to give rise to this curve

so if you look below

each of these

values

youll see a value in brackets so for

the clearance

0

for the volume 0 and for the k

well those are a lot closer exception of

volume

these

are the averages

of the individual estimates that the

population model returned

so conditioning

on the population model that was

estimated

what are the most likely parameters for

each individual and then taking those

parameters and taking their average

we get the values that were simulated

from

so it suggests that maybe this small

simulation set is so small that actually

there are some

maybe nonnormal

distribution aspects to it

so that that posterior mode the maximum

likelihood for those parameters doesnt

match

the average of those parameters

okay

similarly for the between subject

variability

it was a simulated value of the true

value was

for clearance

the nonlinear mix effects

approach retrieved about

relatively close

perhaps for the volume this was almost

and we got

and for ka it was 0 we got 0

again these are the values that come

back from the maximum likelihood

right and so these values are a mode

representing mode if we take the average

if we take this

if we take the standard deviation

of those individual parameters like you

would in a standard twostage approach

that were returned from the nonlinear

mix effects approach you can see that we

now retrieve

those values so its important to think

about how are you summarizing

the population what is an appropriate

statistic that describes

the central tendency and the variability

of that population

it also suggests that maybe with a

sample of five individuals that we

simulated but that is perhaps not

sufficient to get a

or to to to to explore the distributions

if indeed

the modes should match up

id like to take a step back again and

think about parameter distributions just

weve talked about distributions and you

might hear the term density or between

subject variability

and

we really want to think of these things

as as as synonyms

and as reflecting

variability

in physiological spaces

and

these physiological parameters these

physiological spaces that were

interested in

have constraints with respect to what

values these these parameters can assume

and they have to be nonnegative

negative values are not possible so

when were thinking about the types of

distributions we either have to use a

constrained normal because as soon as

you use a normal distribution you can

select a value from minus infinity to

plus infinity may not be very likely but

its still possible and it could give

rise to a parameter that has a negative

value which is physiologically

impossible in the context of our

basic pharmacokinetic models

so the constrained normal is one

strategy or a log normal distribution

so

we can use a log normal approximation so

a natural log of a parameter

uh may

give us

effectively

a gaussian distribution the log of that

parameter plus eta because we can now

select from minus infinity to plus

infinity but were constrained that runs

up against

zero

okay

its typically denoted

as some population value with an

exponentiated inter individual

variability

so this is approximately

log normally distributed

so if we look at

this

graph in the bottom left here what we

have our frequency and probability on

the y axis and a parameter value in the

population you can see that this is a

skewed distribution has a long right

tail this is a log normal distribution

okay

and if we

uh excuse me

if we take the log of these values

what we get

is a gaussian distribution

you can see this upper right curve here

the probability in the y axis and the

prime distribution here

again

if we illustrate it

up above

clearance is equal to theta plus

eta population value of clearance

and the individual sample from a normal

distribution and if this is in the log

domain

we have a

log normal distribution that has been

transformed

so

we just have to remember to back back

trans back transform this okay

this log normal distribution can be

represented as i showed on the previous

slide as an exponentiated function

okay and this will give you a constraint

so you cannot select negative values or

the algorithm will not select negative

values for things like absorption rate

clearance and volume of distribution

and in this case

clearance arises from a log normal

distribution

centered at our population average value

with central tendency omega squared

and that omega squared

is going to be approximately equal in

the log domain

to the square of the coefficient of

variation

so weve talked a lot about

variabilities but one of the one of the

major

advantages

to

applying population pharmacokinetic

techniques and

applying this

hierarchical

pharmacostatistical modeling approach

is to start exploring

explanatory variables what may

describe

this variability so we started saying

well there appears to be variability

between individuals in a population or

between occasions within an individual

and were saying that these are

stochastic processes theyre random

processes

but they may not be strictly random this

is maybe a bin of convenience initially

and we can start exploring what might

explain

some of this quote unquote randomness

what

maybe systematically contributing to

differences in our observed

parameter values and therefore our

observed concentration time profiles or

observed concentration measurements

so why is covariate model building done

well we can use it to help identify

those subgroups of patients especially

those that might be a potential risk of

toxicity or a subtherapeutic effect

perhaps to confirm the absence of an

important of important influence from a

covariate

to increase the mechanistic

interpretability of the model

to understand the trial characteristics

perhaps to generate new hypotheses

perhaps you get a cover relationship

that was unexpected

you say well now we need to confirm this

we need to we need to design a study to

really test whether or not this

relationship exists

we may would want to increase the

predictive performance of the model

and we may also want to increase the

understanding of the studied system

so what are these covariates i think

this is quite perhaps obvious on its

face but lets step through a few of

these

so these could be things such as

demographics

things like gender age size metrics

weight height bsa bmi maybe race

lab values things like serum creatinine

bilirubin albumin phenotypes and

genotypes

disease parameters perhaps the baseline

status of disease the severity of an

injury

the etiology of the disease maybe you

understand something about how the

disease evolves and might affect other

systems and you want to you want to

evaluate this

could be therapy related perhaps

a person is getting dialysis

or is going for surgery and is going to

be on a bypass machine or theyre taking

concomitant medications that could be

interacting

either increasing or decreasing the

concentrations of a drug of of of

interest

there could be habits or environmental

factors

uh diet

smoking alcohol intake time of day or

or or time of year

maybe study related maybe certain sites

seem to have differences you can begin

to explore these things

once you establish

these hierarchical bins of variability

so types of covariance that we may

want to incorporate so

there are different sort of types of

data that youre going to get in there

so there are different ways of actually

incorporating or implementing coverage

analysis based on this

these include

continuous variables things like age

dichotomous variables perhaps a

bivariate

variable such as sex

there could be ordered categorical

values so ordered categorical means

there are categories but those

categories have

a hierarchy they have an order

so things like uh whether or not youre

a poor metabolizer an extensive

metabolizer an intermediate metabolizer

theres an order to this

and

you you even though theyre separated by

by categories you wouldnt incorporate

that as a continuous variable

theyre perhaps nonordered categorical

things like

perhaps race categories

that

we dont maybe have a specific

understanding of

whether or not there is a systematic

ranking of of these factors but there

are multiple factors that need to be

evaluated individually

so were building these covariate models

and we could actually add the covariates

on our

fixed effects on our structural

parameters how does this affect

clearance

how does this affect

absorption rate how does this affect

volume of distribution

we may also consider that the

interindividual variability in addition

could be affected maybe within one of

these subgroups

the between subject variability is much

smaller for example if you have a poor

metabolizer genotype

perhaps the variability in the

metabolism is

very very low because they have almost

no metabolism of a particular drug

so driven to sort of a maximum condition

okay

maybe it affects the interoccasion

variability or the residual variability

maybe there are interactions amongst

covariance maybe there are multiple

interactions to be considered

simultaneously were going to focus

mainly on the on on the implementation

the structural para on the structural

parameters on the fixed effects things

like clearance volume of distribution

and absorption rate

so there are some dangers though with

coverite analysis as well

if we maybe consider too many covariates

you may end up defining what are false

or irrelevant

covert relationships

that actually will be

detrimental to the predictive

performance of the model

maybe leads to the collection of

unnecessary information maybe this

information or these data are somewhat

intrusive or burdensome to your patients

to gather

perhaps you will generate hypotheses

that result from fro from this that

actually

dont bear fruit with respect to

interpretable results

okay

or with respect to

newer novel new or novel findings

because the false positive rate

is so high in your in your approach to

covariate modeling

and if these are disseminated therell

be less trust

in the identification of relationships

when they are indeed

a true relationship

so in terms of some general steps in

cover model building

you start of course with all coverage

parameters of interest

but its important to maybe think about

the scientific

plausibility

this is you want to be careful about

this as well because you dont want to

maybe look at things that you already

know exist maybe there may be some other

factors that that that could be

important that you may miss if you maybe

narrow this too too much

but thats up to the investigator to

decide

then maybe setting a a a reasonable

threshold for statistical significance

so maybe you

want to do some learning and so youll

set a relatively

uh relatively low

threshold for statistical significance

but then refine this with a much more

stringent or much more conservative

alpha value for example

a very important component is to say

well were not only interested in

statistical significance but how

clinically relevant is this covariate

how much does it change the clearance

the absorption rate the volume of

distribution

does it make a difference in the overall

exposure in a patient

does it make

a difference in a sufficient proportion

of the study population

to be of concern

and then feed this all forward into our

final coverage parameter relationship

so just to think about the

parameterization of covariates you may

take

a factor such as age

and you may discretize it here we have

the central circle here we have

clearance on the

on the yaxis

and here we have two bins less than

years of age and greater than years

of age and you can see that theres a

higher clearance for our younger

study subjects than our older study

subjects

maybe thats

the the the the the way you would like

to to to represent the information maybe

the way the studies were designed you

dont have more detailed information to

get at maybe a a more information rich

relationship

this discretizing the covariate actually

removes quite a lot of the information

perhaps actually theres a linear

relationship with age and we should be

concerned about a continuously

decreasing elimination of the drug over

a wide age range

perhaps theres a nonlinear

relationship with age maybe the

clearance stays relatively stable until

some particular age threshold then

appears to decline

its important to evaluate your data

as well to make sure you have a

sufficient signal range in your

covariate value

so that you could actually begin to

identify

maybe and discriminate amongst these

types of relationships and therefore how

you would like to parameterize that

model relationship in your population

pharmacokinetic model

so one can use exploratory data analysis

this is a matrix plot

and here we have

von villa brands factor body weight

bmi and age and you can see along this

this

diagonal of this matrix we actually have

the the distribution of each of these

parameters so you get a sense of the

range

the distribution the central tendency of

each individual parameter

and then we have a plot of each of these

values heres one villa brand versus

body weight here is a very nice body

weight versus bmi so youd expect a very

strong correlation

you can see that the correlation

coefficient is 0 so these types of

these types of figures are helpful

in

evaluating potential covariate

relationships and understanding

what the range of covariate values are

in your data set and how they are

distributed

so whether or not you likely have the

kind of signal range to effectively

identify a covariate relationship

it also

is important to understand what the

correlation is amongst your covariates

because you may not want to incorporate

multiple correlated covariates as

correlated covariates will contain very

similar information there may be some

exceptions to that but its

perhaps going to to to to it is

something to be to be cautious uh

cautious and conservative about

so why do we look at covert correlations

i think we

we touched on the fact that the

correlated covariates partially carry

the same information

they may also carry

uh somewhat different information too it

may be that theres a correlation but

theres individuals where that

correlation is broken

that are have a very unique response and

maybe thats thats important

information but that something has to be

evaluated very carefully

so

one of the one of the key issues is that

if you have simultaneously

incorporated correlated covariates as

fixed effects

you may get an increase in parameter and

precision okay you may increase the

instability of the model and you may

even get counterintuitive models

because of the interaction of these

covariants

you may also decrease

the predictive value

if you have too many of these these

correlate or if you incorporate

correlated correlated values okay

but its really up to the modeler to

make the to to to make the uh to to make

the the determination of which

correlated covariate carries the most

predictive value

okay

and this is typically evaluated using

things such as crossvalidation or

external validation of the model thats

probably the most

uh sensitive to these types of

overfitting issues

along with epsilon shrinkage that well

talk about in a few slides

so identifying covariates weve kind of

talked about this very briefly we looked

at sort of covariates amongst themselves

but then we may want to look at

our our individual estimates

or are residuals even in our predictions

versus covariates

to see whether or not there appears to

be a systematic pattern a systematic

relationship

a nonrandom pattern

that appears to be

present

between a specific patient

characteristic or covariate

and

the model fit or the model deviation

from the fit

unfortunately this is often done once

youve established all these

hierarchies

of

of of of variants and have have that

initial population model established

again we need to be

cognizant of preventing false

cover parameter relationships probably

never be fully immunized to this

but again considering the scientific

plausibility

the clinical relevance whats the impact

on my patient

whats the impact

on

the specific parameter and on the

outcome of interest how much does this

change the concentration time profile

the overall exposure

even at the maximum lets say the

maximum possible deviation for a given

coverage relationship

you may want to use a

more stringent

uh alpha value

demand higher lower p values

you want to explore covariates for which

sufficient variability exists in the

study population

so that you do not have

one or two maybe

influential individuals with an extreme

value of a covariate driving the

analysis

and also if you can roll this out into

multiple studies and see

a confirmation

of the covariate relationship that is

often

something that can give you greater

confidence in the

covert relationship the correct

relationship remains identifiable

with the proviso the studies youre

evaluating them are adequately powered

and adequately designed to to to

identify such an effect if the effect is

present

theres some potential issues

additional issues with building

covariate models

if weve got the structural model wrong

then the covariate may influence the

wrong parameter

if an existing covariate parameters

lacking in the model that coverage is

likely to be significant if tested on

another parameter so you might depending

on the order in which you include the

covariate relationship you might

identify

a covariate on a parameter its

correlated with another one but its not

the true relationship and in two slides

actually ill show you

an example

now you may not have it between subject

variability

in your model on a particular parameter

perhaps it wasnt estimable when you

were developing that population model

you can still incorporate covariates on

those parameters but theres likely to

be a high false positive rate

one thing to really keep in mind is

whether or not

you have reason to believe or there is

there are evidence that suggests that

that the cover relationship may be

highly nonnormally distributed

if its highly nonnormally distributed

lets say its multimodal or has some

very different distributional

characteristic

the between subject variability wont be

estimated because it has to be made so

large to encompass this variability

that the algorithm trades off poor

prediction of a few data points

for

select for for

for selecting a smaller between subject

variability or nonexistent between

subject variability on that particular

parameter

and so these these can be explored ive

given a few references the first here is

a classic this is janet wades paper

from 99

and really illustrates the sensitivity

uh of

the

selection of the covariates based on

whether or not you have that structural

model correct thats not on your mixfx

model correct

the savage and carlson

paper and aaps journal and vijay

ivaturis presentation here from page

address issues of shrinkage and covert

identification

so

when you add the coverage does it

improve our diagnostics or basic

goodness of fit maybe empirical based

estimate based diagnostics maybe visual

predictive checks

uh is the value of the relationship

realistic is it well determined are

those parameter estimates robust

does the unexplained variability

decrease and does the objective function

value decrease

is it driven by

influential individuals so i would

encourage everyone to think about these

things as they move through their

modeling process

i mentioned the inspection of the

empirical base estimate so here we have

the normalized parameter deviation 0

versus creatinine clearance

okay you can see a relatively strong

relationship in the left panel

and somewhat of a relationship in the

right panel this is

the original relationship and this is

what happens if you include the

a different

covariate like body weight and age

perhaps your relationship becomes less

strong because the creatine clearance

actually contains that information its

correlated to those elements

so we want to be cognizant of

incorporating these things this gets to

the scientific plausibility and what may

be contributing to

observed variability in your models

so another

diagnostic thats used

that helps us to

understand when we need to do additional

model evaluation with respect to in

particular covariate modeling

is to evaluate something called

shrinkage in the model

and say well what is shrinkage that

sounds sort of strange

well

remember that when you estimate the

population model using a nonlinear mixed

effects or a bayesian approach you get a

central tendency and you get a

variability

and then conditional on each

individuals data

you get

values for that individuals clearance

volume lets say an absorption rate if

we use our basic example

okay

if the data for each individual are

noisy or very sparse or not very well

collected

the algorithm is going to say well i

dont have a lot of information to

justify

picking

an individual value thats very

different from that population value

from the population central tendency

okay

and so the values that are selected for

those individuals are going to shrink

towards the central tendency theyre

going to pick things that are most

likely at the population level because

their data are not sufficient

to

to justify

selecting an individual ada random

effect to change their

parameter values

and so this becomes apparent

as calculated in this equation here

where we look at the standard deviation

based on those individual estimates

versus what the algorithm told us that

population variability should be

and youll see that if

these match

shrinkage is low this this equation goes

to zero

whether you use the standard deviation

of the variance

if

these values are being chosen for the

individuals close to the central

tendency this ratio

will be very small

the shrinkage will be close to one

or a hundred percent usually often

reported as a percentage

and therefore

you need to be cautious when youre

looking at empirical bayes estimates

because you may miss cover relationships

because those values are selected close

to the central tendency

so you need additional

model evaluation techniques and you

should also be cautious about the values

that you do identify or the covert

relations if you do identify

at the level of the residual error we

get a different type of shrinkage its

called epsilon shrinkage and you can see

that its one minus the standard

deviation of the weight of the

individual weighted residual or the

variance of the individual weighted

residual so what does that mean it means

if these weighted residuals are really

small the shrinkage is going to be close

to one

and youre saying okay

so why should i be concerned about that

well it means that your model is almost

perfectly predicting your data which is

not particularly likely

and you need to be concerned that you

may be overfitting you may have too many

parameters in the model or too many

covariate relationships identified

and yes youre getting virtually perfect

perfect predictions of the data you

observe but when you go to predict into

the next data set or the next individual

your predictive performance may be very

poor because the model is so highly over

fit

and so this typically when you have high

shrinkage you want to do additional

simulation based evaluation

and ill leave it to you to look at

visual predictive checks and normalize

predictive distribution error checks

theres actually an excellent paper that

just came out in cpt psp a white paper

from france montreals group that i

would recommend

folks look into i believe its just a

few months old now

i think weve talked a lot about the

consequence the consequences of ada and

epsilon shrinkage

so

for our empirical bayes estimates in

addition to shrinkage you may also look

at a nonnormal distribution so remember

we said the distribution should be

centered at zero with variance omega

squared well what if its not centered

at zero

so maybe its suggesting the underlying

data distribution is nonnormal for

those ebes

even though the assumption for the

algorithm is that it needs to be

centered at zero

okay

the mean value for those ebs may be

significantly different from zero even

for a correctly specified model if

youve got a situation with significant

shrinkage

you may actually end up with

correlations that youll either thatll

either not be discovered or maybe

spurious

i think we mentioned that you know i

think it was already mentioned that

epsilon shrinkage is more sensitive to

overfitting

and your weighted residual shrink to

zero

and your model evaluation therefore with

your weighted residuals becomes less

effective

because your models over fit and your

deviations between your observations

your predictions

are

very very small

so weve talked a lot about covariates

and incorporation of coverage but what

are the what is the main approach to

doing this

the main approach is forward known as

forward addition and backward

elimination

so you start with no coverites in the

model we talked about this you test each

covariate individually you determine

which one improves the model most

significantly and then you add this

covariate to the model

you repeat this until there are no

significant covariates remaining

then you have this full model

then you start deleting

covariates

and you can go in the reverse direction

or you can start with all the covariates

in the model and you can delete each

covariate and look at whether or not the

model gets worse

so there are multiple approaches theres

a schematic in the next slide i think

that illustrates this

nicely

the forward addition you start with a

base model with no covariates you

evaluate your diagnostic information

you test all the suspected covariates

adding them one at a time

you pick the model with the most

significant covariate

okay you have a new covariate based

model then you start adding covariates

one at a time

if it further improves the model you

keep it if it doesnt you leave it out

eventually

you have added all the covariates and

you end up with a model that only has

the ones that made significant

improvements to the model

and then you start removing the

covariates in a different order

okay one at a time

does the model get worse

if it doesnt it if it if if it if it

doesnt get worse then you remove that

covariate and you repeat until

any elimination from that model results

in a worse model statistically

and you end up with your covariate model

resulting from backward elimination

so just sort of to recapitulate and sort

of summarize with respect to individual

population information

when were determining population

characteristics for pharmacokinetics or

for actually for anything biomarkers

whatever youre measuring

if you have a lot of individuals and a

lot of data per subject then youve got

very robust individual population

information

if we dont have

many individuals but we have a lot of

data per individual then our individual

information is most robust

if we dont have

many

observations within an individual we

have lots of individuals then we can

come up with a reasonable population

model and make inferences on the basis

of that model

if we

uh have very few individuals with very

few measurements well thats i think

pretty obvious on its face that were

not going to get very far in terms of

being able to make inferences

so to

to further recapitulate population pk

modeling we have a model that comprises

a structural or a structural model part

of the fixed effects our clearance our

volume rka

that structural model is interacting

with covariates

okay and these covariates can represent

uh can be individual specific

characteristics

that modify the structure and then we

have a statistical model that describes

variability again the structural model

these are the functions describing our

typical concentration time course this

is our one compartment two compartment

three compartment model or pvpk model

whatever model youre using

we represent these often as differential

equations

this could be for example the one

compartment model following an iv bolus

dose

we have our statistical model that tells

us about variability around that

structural model at the between subject

level at the between occasion level and

at the residual unexplained level

okay and again between a subject between

occasion residual

and then we have

patientspecific characteristics

that explain variability ie the

variable is predicted by

subjectspecific characteristics and

thats our covariant model

so we encompass these three key

components here

for example things like weight age and

genotype

so in conclusion our population

pharmacokinetic approach is ready to

capture both group and individual

characteristics

central tendency variability and then we

can start to understand the individual

characteristics conditioned on their

data

it can be used for simulation with

variability and uncertainty around

parameter estimates

and its

very useful in situations where only

sparse data are available you can

imagine for special patient populations

etc this is particularly useful for

example in the icu setting perhaps in

geriatrics in pediatrics

and with that i will draw this lecture

to a close

im happy to address queries and i

believe there will be a

supplementary set of slides with some

with some specific examples related to

population pk model provided

as part of the slide set thank you