hello Im Rob bees and Im faculty at

Buffalo and uh Im uh presenting today

on a

complimentary uh vignette to the lecture

on population

pharmacokinetics uh these are broken

into multiple parts and this is part one

part one focus on exploratory data

analysis and the initial elements that

are necessary prior to to undertaking a

formal population pharmacokinetic

analysis so theres some general steps

in a population pharmacokinetic modeling

project the first of course is to

collect the data so some design some

experiment is designed or question is is

is arrived at an experiment is designed

data are generated and then we collect

the

data uh we need you know one needs to

understand exactly how the data were

generated uh theres data formatting

involved depending on what software or

what approach you might use for the

analysis uh theres an exploratory data

analysis step and then model build
approaches a general schematic is shown
in this slide with the general steps for
the population pharmokinetic modeling
project and our part one of our vignette
really focuses on this top part here the
research question experiments data and
visualization data of data with a
particular focus on the nature of the
data and the visualization of the data
before we get to things such as model
context to to explore that will be taken
up in the
subsequent uh
vignette again were going to use
actually a case study with copine and
these are kipine concentration
measurements that were used for
population pharmacokinetic analysis for
data that were generated from the
clinical antic psychotic Trials of
intervention Effectiveness this is a an
imh funded study that
completed
uh over over a decade
ago um were going to again familiar

ourselves with the study design and
understand the data generating process
were going through going to go through
a very basic uh illustration of how to
construct a data set for population
pharmacokinetic analysis in this case a
nonmem ready data set nonmem is a one
of the many software available to
conduct this type of
analysis uh and then uh uh we will go
through an exploratory data analysis
evaluating the different characteristics
of the data that were
collected uh this is done prior to the
formal model building process and is
often done with a data wizard uh or with
a statistical programming language such
as
R uh things that we will be evaluating
are the numbers of observations numbers
of individ in this case concentration
observations
uh number of individuals what were the
doses administered what were what are
the patient specific
characteristics what is the distribution

of times after dose for the
concentration measurements and then any
information that one can glean from a
typical population pharmacokinetic study
with respect to the concentration versus
time profiles both at the individual and
aggregate
levels uh this uh is important in terms
of evaluating uh those profiles at the
level of patient specific or study
specific
characteristics another part part two of
the
vignette uh will cover the population
pharmacokinetic model building and
Analysis strategy and is outside the
scope of this initial
vignette so a key step is the
familiarization with the data generating
process or the study
design when considering the study design
one should evaluate whether or not this
was an observational or a randomized
study this will
impact uh the consideration of potential
confounders was a study conducted in a

highly controlled setting like a

clinical research center what was the

concentration sampling Paradigm was it

random was it

optimized what would what was the dosage

Paradigm what was the frequency of doses

what were the range of doses

what covariants were collected IE what

patient specific characteristics or

demographics do we have in the data

set and then what was the nature of the

sample collection were biomarkers

collected in addition to concentration

was the timing recorded

appropriately so lets move on to the

construction of the data

set so when one is assembling data from

a study such as Katie study one may get

data in

multiple uh table formats here I have a

demographics

table uh showing weight um and sex by

study

subject I have a table of doses so you

have the study subject you see there are

multiple doses listed and the time of

those doses in

hours and then I have a table of

observations so here we have the study

subject the times the plasma

concentration

measured so you notice the study subject

is the linking variable

here and for a population analysis the

study subject that is the that is the

basic level of analysis one considers

that as sort of the basic unit of

analysis respect to the determination of

parameters

and so if we look at a typical nonmem

data set it is actually assembled from

those components you see the study

subject ID appears at the far

left and uh different uh elements here

but we will go through these uh

individually in a

moment Now sort of superimposed the

original table formats or at least

excerpts or subsets of those initial

table formats to illustrate this

see that our table that indicate study

subject and the dosage histories are

captured in the overall analysis data

set as the study subject ID here labeled

pound ID and you can see these are

listed

here the dose in milligrams in this case

appears in the AMT or amount column and

you can see that they are noncontiguous

rows for most part because there are

observations down in

between and so this is now

sequenced by

time similarly the next table showing

steady subject time and plasma

concentration appear again in the ID

column time and the DV which in an nonm

parament stands for dependent variables

this this is your observation this is

your

concentration and you can see now that

weve got weve got weve got plasma

concentrations interspersed here so the

nonm data set runs sequentially so you

go you have a

dose and you have an observation and it

runs sequentially in time time must be

ascending the other thing to note is

that we dont have any less than symbols

and we dont have the negative timing

typically you do not want to avoid

putting in a negative time off the newer

versions of nonm can manage that U and

you only

numerical uh elements can be present in

the data so you you cannot have a less

than symbol you cannot have Alpha

Characters uh in the data set thats

mainly because this software uses

Fortran but even other analytical

software for the numerical components

will require numerical

inputs finally we have the demographic

table so again study subject right so

this is indexing

here weight you see weight appears and

weight is numerical and then we

have sex sex is listed in the original

table as female and

male but has to be

converted to a numerical flag in order

for this to be incorporated into the

analysis and so you see we have a column

we call were calling sex

M uh and zero is female one is male and

you can see these are the first two

individuals are those is female and

these

values are noted on every single Row for

those

individuals so now that weve put

together a general uh data

set there are a couple of other

components here that were

Incorporated these are the missing

dependent variable and the vent ID and

these are nonmem specific indicators

and Beyond the

scope uh of of the presentation

today but I will just make a quick note

that at least the missing dependent

variable is used to indicate when

theres an input process when theres

not an observation so here when you give

a dose you have a missing dependent

variable because you cannot have a dose

and an observation an input and an

output in the same

row um and so youll notice the missing

dependent variable is zero would you

have a plasma concentration this is

Illustrated

here so lets move on to the exploratory

data analysis component we thought about

the study

design we formatted the data set so

were ready to start potentially doing

some some

analyses and now were going to do an

exploratory data

analysis evaluating the information we

have in this data set prior to

modeling were going to consider the

available observations and demographics

the covariate distributions the

correlations the signal ranges how

broader those signal ranges are they

something we want to test because we

have enough variety enough heterogeneity

in those measures to be able to pick up

a difference if there is

one going to evaluate the concentration

measurements the units the range of

observed concentrations any correlations

with covariates doses administered the

roots of administration and

administration rates

Etc this is also the opportunity to

evaluate extreme values in the

observations when one does expor data

analysis in particular

visualizations this these become

apparent one we also want to evaluate

the proportion of measurements that fall

below the limit of

quantitation a determination of samples

that are missing or damaged so these are

two separate components below the limit

of quantitation means a measure

measurement was made but the limit of

the quantitation of the assay was um

crossed and

therefore uh the measurement

value uh may be reported just as below

the limit of quantitation or a number

that does not conform to the stringency

of the assay as it was developed for

quantifying those

values again a preliminary evaluation of

observations with an evaluation of

outliers potentially evaluating the

concentration time profile if we have

sufficient sampling

a preliminary determination of the

potential of the structural models to

test and the preliminary calculation of

potential parameter values overall

clearance for example the volume

distribution to use the starting values

all population PK analyses are nonlinear

regression problems and you have to

start with initial

values initial

estimates so for this exploratory data

analysis as I mentioned we have the

copine concentration measurements in the

controlled antic psychotics trial of

intervention Effectiveness Katie this

was a treatment study with patients in a

schizophrenia arm and an Alzheimers

disease arm and sparse concentration

measurements were taken had approxim

just over two concentration samples per

individual so as we explore the data

some tabular output in terms of the

population

description so number of

observations in the schizophrenia arm we

concentration observations and the

Alzheimers are  so 9 observations

all

told the bottom table shows a

distribution by selfreported

brace uh and uh so you can evaluate the

relative contributions here and whether

or

not they appear to be uh representative

of a population at

large uh one thing to keep in mind as we

go through these different ele elements

is whether or not there appear to be

systematic differences between the arms

of the study or even amongst the

covariants U this may potentially

confound identifying a specific effect

if there are strong

correlations similarly here we have a

table top showing the distribution uh BX

and the schizophrenia and the

Alzheimers arms we had

0 uh individuals and weve got  in

the schizophrenia arm 9 in the

Alzheimers arm uh in the schizophrenia

arm it was predominantly male almost

of the of the individual patients the

Alzheimers uh study was more evenly

split with actually slight majority of

female

patients in terms of smokers and

nonsmokers see that uh we had a

majority of smokers in the schizophrenia

arm

uh versus the Alzheimers arm so this

could be an issue because it was an

effect of smoking you might have smoking

and and arm confounded because of the

different distributions so this is

something to keep in mind when youre

exploring

explanatory

variables in terms of the weight

distribution you can see as well the

schizophrenia arm the weights

were substantially

higher uh similar ly for the age

distribution in the bottom table we have

uh Alzheimers course arm the age was

substantially higher not surprising with

respect to the type of study so again if

youre interested in in if you detect an

age Effect one has to consider that

there may be a different disease process

as well these are different arms of the

study dosages were different this is

these are all elements to keep in

mind so while tabular approaches are

helpful it also is important to

visualize the

information so this next slide which was

created using R and a package called

ggal shows a matrix plot for patient

demographics this is for both the

Alzheimer and schizophrenia studies

combined so we you see on the diagonal

of this

Matrix are the uh distribution for

example for age

so you can see it looks quite

bodal uh the distribution for weight

its a bit of a shoulder here so maybe

not entirely glossi in you see also the

distribution by

sex female and

male uh and you see the distribution of

smokers in the combined studies you see

thats almost 00 you combine the

studies in the off diagonal elements

were able to

evaluate

correlations and other aspects of the

distribution so you see a

relationship with sex and age across the

studies smoking

status and age as

well uh and similarly await distribution

difference for uh sex with males being

SL the

heavier uh smokers and weight do not

show at least for the combin set this

element in terms of the other off

diagonal elements here the size of these

squares represents the number of

individuals here we have nonsmokers and

smokers by

sex so males and and

females okay and then in the bottom left

set of off diagonal elements we actually

see the individual histogram so you can

see the distribution of values for

example so for age BX okay for

weight right

BX okay you can see the shift here this

is these are the males these are females

males typically being a bit heavier but

you can get a sense of signal

ranges now one of the major components

is that we have these we have these

subcategories so what happens when we

subset these characteristics lets say

by

by by by sex and smoking thats whats

shown in this next set of

plots and again on the diagonal we have

the distributions but this is now by

different sex groups you can see that we

have more females in the older group but

in the younger

group the females and males are pretty

much

matched similarly for

weight the males are somewhat

heavier

okay

um and then

by uh smoking

status uh you see

that uh males tend to be more likely to

be

smokes similarly if we subset by smoking

by by smoking and

nonsmoking you can see that most of the

smokers are in the younger age group

weights overlap but we again see this

big difference by uh by sex so many more

male smokers for

example

similarly uh the you again have the

distributions for the individuals and

look at in terms of

correlatives

um we have there is a significant

difference here so if we

consider uh sex and age have nonsmokers

and smokers theres a similar uh Trend

here so we have uh typically younger uh

individuals smoking in this in this

study set um by weight again weight and

you you see that um for smokers and and

nonsmokers in in females as smok a

little bit uh uh light

than in males but it doesnt appear to

be much difference in weight so we start

to get some insights about the

distribution of these

parameters so what if we

consider the study arms

separately so actually going back to the

previous slide you can see that there

was some you know there were some

differences there were some sex and you

know weight differences in sex and

smoking status and that sort of things

so if we go and look at only the

schizophrenia

arm notice now that the age

distribution is quite

different its

less uh it still has a shoulder but its

less sort of

bodal and the uh sex and age

differences are less

pronounced smoking and age differences

are less pronounced individuals tended

to be younger we still have slightly

heavier males and females in this

study um and so that might be important

information is there an effect of the

disease or is it the effect of the age

group or is it effect of the weight if

youre looking at specific effects on

pharmacokinetic parameters at the

population

moment you subset

this by sex

groups or by smoking and nonsmoking

groups see that many of these

differences are not preserved looks

previous previous graph so again age

weight smoking and along this Matrix

here you see the these distributions are

largely overlapping at least in terms of

the continuous

variables uh when you subset along those

status within the

study similarly you can see

some uh weight uh difference sub by

smoking and nonsmoking groups nonsmokers

smokers um Nails actually tend to be a

bit

lighter what about the Alzheimers arm

so here again you see that the age now

has shifted to the

right the weight now again does not have

as much of a shoulder looks more

gossan the uh sex distribution is more

even and the smoking distribution is

quite

dramatically different very few smokers

actually in this uh in this

study and if we subset by the uh sex and

smoking

status you see that for the different uh

sex

groups um you can see that theres

almost

no uh smokers

here for female versus

male and

that we do have a shift in the weights

females are lighter than males in this

case ages though are quite closely cing

so things to keep in mind in terms of

potential confounders again this is done

with G galy and its I think an

important uh evaluation of the data

before we jump into the

[Music]

analysis so those are patient

demographics what about par what about

elements that are related to the pH

cinetics

directly so these tables show Doses and

the frequency by trial arm so the top

table focuses on the dosage interval so

there were hour dose intervals and

hour dose intervals we can see that

in the schizophrenia

study the majority of individuals

receive the doses twice a day  of

those

individuals  receive the doses every

hours

in the Alzheimers arm no patients

receive doses every  hours 00 of

patients receive them every  hours now

this might be of

concern if the drug has a relatively

fast halflife perhaps youll have more

measurements for alumin

quantitation because of the dosage

interval being wider before the patients

are

redosed overall if we combine the

studies about 0 of the patients were

receiving uh kipine every  hours and

were receiving them every

hours some patient idea IDE some

patients had both dose intervals

implemented so they had some periods

when they were on every  hours and

some periods when they were on every

hours what about the doses administered

so this uh bar chart is showing doses

daily doses for the schizophrenia arm

youll see that we have doses that go

from

uh two uh 00 to 00 uh milligrams in

the schizophrenia on not very many low

doses so this would be 00 milligrams

twice a day for example these were all

given every  hours um and a relatively

even dist

distribution uh across the the dosage

ranges so well represented for the

analysis with respect to the Alzheimers

disease arm you see the dose

distribution is quite different its

well given every uh  hours and the

doses R dosages range from  to 00

milligrams daily the doses are listed

here in micrograms that matches our

units in the actual analysis and the

majority of doses here at the low end

milligrams so are we going to see many

more again limited quantitation issues

with the Alzheimers disease on

because theyre lower doses over a wider

interval then again theyre older

patients and maybe they eliminate the

doses more slowly so these are all

things to keep in mind with respect to

uh the population pharmacokinetic

[Music]

analysis another key element to evaluate

is how much information do we have what

is how many concentration measurements

do we have across the entire

concentration time profile we have

absorption processes distribution

processes in elimination processes and

these are sparse samples so its

important to understand whether or not

you have sufficient sampling across the

entire dosage interval so that you can

inform those parameters so on the left

this is from the schizophrenia arm we

have the time since last dose

distribution with a hour dose

interval we have a frequency on the y

axis and the time since last dose on the

xaxis what you see is that we have

quite a few samples shortly after the do

so well get the absorption processes

reasonably well but as we get closer to

the end of the interval we have fewer

samples less

information with respect to the time

since last dose distribution for the

hour dosage interval you can see we

again have some we have less information

actually in the absorption profile and

then we have more information in the

tail phase of the distribution V

concentration time

profiles for the Alzheimers disease

study we see a similarly bodal

distribution which probably relates to

clinics being closed in the middle of

the night and patients coming in for

usual clinic hours and depending on

whether or not they take their their

dose in the morning or the

evening again time since last dose we

have the frequency on the y axis time

since last dose on the x axis you see

that we have information that will that

will help understand the absorption

process absorption profile

and then the tail distribution as we uh

uh go towards the end of the dosage

interval or reasonably well uh reasonbly

well

captured so what about actual

concentration measurements well heres a

plot with the uh dose normalized

concentration remember we had multiple

concentrations this is for the

schizophrenia

arm and each of these black lines

represents uh

an individual with con with the time

after time since last dose on the xaxis

here time since last dose and theyre

connected so you can see profiles for

every

individual the green line in the Shaded

area is basically a

smooth

through these profiles so you can see

that

overall we capture a the up slope and

down slope of the concentration time

profile that one would expect with this

uh with this drug and we can evaluate

whether or not there appears to be

multiexponential character we can also

evaluate whether or not this may be

sufficient to actually allow us to

calculate some preliminary initial

estimates for clearance and volum

distribution for

example now there were multiple doses

given though and maybe theres a

nonlinearity in dose to evaluate that

again we have dose normalized

concentration on the y axis and time

since last dose on the xaxis but now we

have a color coding for the different

doses again for within individuals we

have connectors connecting the

doses and blue the light blue are the

highest doses the black are the lowest

Doses and you can see that these pretty

much overlap so if you normalize the

concentrations it suggests that linear

processes are are are are driving the

pharmacokinetic

profile

another evaluation is to look at the

individual patients time concentration

profile so this Matrix of plots is

showing concentration on the y axis

times the last do on the x axis again

this is a schizophrenia study for for

the hour interval and this is useful

for identifying whether or not there

appear to be any significant

outliers uh by individual and we dont

really see any here similarly we look at

the hour interval

you see this is for the schizophrenia

arm and again dont necessarily see any

outliers there a few maybe maybe higher

concentrations here that we want to

investigate uh but generally

speaking uh

reasonable for the Alzheimers disease

arm here a sample of patients again this

is a matrix plot why each panel

represents a different individual

y axis is the dose normalized

concentration and the xaxis is the time

since last dose and again you can use

this to inspect for potential

outliers so in conclusion a complete

exploratory data analysis is critical

prior to beginning the actual model

building process for a population from

kinetic

analysis insights on what types of model

structures can be supported what

covariants may be considered and what

potential confounds can be

preidentified

and so our now that weve done the

exploration of the

data we can begin the population

pharmacokinetic analysis and this will

be covered in part in the vignette part

two Id also like to

acknowledge uh two individuals who

contributed to this uh Christen Bigos is

faculty at Johns Hopkins

University uh she

facilitated access and provided subal

analysis on

copine uh data

set and uh an MS candidate Master

student at the University of Buffalo har

on gal we prepared many of the graphics

that I that I showed using those R

packages thank you very

much