Combining Studies Involving Repeated Drug Administration
… Population Analysis & Automation …

Ray Boston

December 10, 2020

For Course Participants … Help and Explanations

Please contact me at drrayboston@yahoo.com

This study

Four volunteers were given an **IV bolus** dose of **10 mg** of compound A. Assume a body **weight** of 70 kg. In the first part of the study, plasma samples were obtained from two **males** (cf the total of 4 volunteers and the concentration time data were as plotted below . What do we **learn** from these plots?

Do these two subjects have different sizes of **volumes** of distribution? Are their fractional **clearance** rates the same?

```
scatter qc qo time if c==1 & study<=2, c(l .) ms(i oh i) yscale(log) ///
  by(study, style(compact) ti(Semi Log Plot of Responses)) lc(red .) ///
  mcolor(. green) ylabel(100 300 1000,angle(0)) legend(row(1) ///
  lab(1 "Subj") lab(2 "Obs"))
```


```
scatter qc qo time if c==1 & study<=2, c(L .) sort(study time) ///
  ms(i oh i) yscale(log)  ti(Semi Log Plot of Responses) lc(red .) ///
  mcolor(. green) ylabel(100 300 1000,angle(0)) legend(row(1) ///
  lab(1 "Subj") lab(2 "Obs"))
```

By **reading** the population studies tmp file (data.stmp) using Stata's '**insheet**' command we are able to produce the plots on the prior slide

```
cd "c:\stata2019\New Biostatistics Program for Penn Vet aug 12 2020\"
cd "demonstrations\multiple studies 1

insheet using data.stmp, clear
describe
scatter qc qo time if c==1 & study<=2, c(l .) ms(i oh i) yscale(log) ///
  by(study, style(compact) ti(Semi Log Plot of Responses)) lc(red .) ///
  mcolor(. green) ylabel(100 300 1000,angle(0)) legend(row(1) ///
  lab(1 "Subj") lab(2 "Obs"))
more
scatter qc qo time if c==1 & study<=2, c(L .) sort(study time) ///
  ms(i oh i) yscale(log)  ti(Semi Log Plot of Responses) lc(red .) ///
  mcolor(. green) ylabel(100 300 1000,angle(0)) legend(row(1) ///
  lab(1 "Subj") lab(2 "Obs"))
table study c, row col
```

```
. describe

Contains data
  obs:            68
 vars:            13
-------------------------------------------------------------
              storage   display    value
variable name   type    format     label      variable label
-------------------------------------------------------------
label           str7    %9s                   Label
study           byte    %8.0g                 Study
studyname       str12   %12s                  Study Name
c               byte    %8.0g                 C
category        str5    %9s                   Category
tc              byte    %8.0g                 TC
time            int     %8.0g                 Time
qc              float   %9.0g                 QC
qo              int     %8.0g                 QO
sd              byte    %8.0g                 SD
fsd             float   %9.0g                 FSD
theta           byte    %8.0g                 Theta
weight          byte    %8.0g                 Weight
-------------------------------------------------------------
```

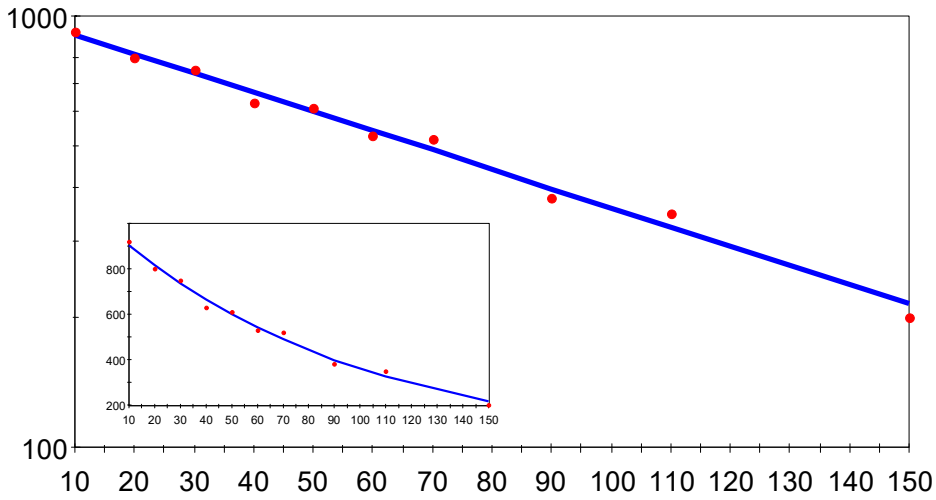
```
. table study c, row col

------------------------------
           |          C
     Study |     0      1  Total
-----------+------------------
         1 |     7     10     17
         2 |     7     10     17
         3 |     7     10     17
         4 |     7     10     17
           |
     Total |    28     40     68
------------------------------
```

```
A SAAM31                                    Subject_1
2       10
H PAR
C
c# K1=K(1)
c# L01=L(0,1)
c# p1_Vd=p(1)
c# p2_Cl=p(2)
c# p3_Cmax=p(3)
c# p4_t1/2=p(4)
c# p5_AUC=p(5)
c# p6_AUMC=p(6)
c# p7_MRT=p(7)
C
   K(1)      1.002162E-01  0.000000E+00   1.000000E+04
   L(0,1)    1.025612E-02  0.000000E+00   1.000000E+04
   ic(1)    10000
 p(1)=1/k(1)
 p(2)=p(1)*l(0,1)
 p(3)=ic(1)*k(1)
 p(4)=alog(2)/l(0,1)
 p(5)=k(1)*ic(1)/l(0,1)
 p(6)=k(1)*ic(1)/l(0,1)/l(0,1)
 p(7)=p(6)/p(5)
H DAT
   p(1)
   p(2)
   p(3)
   p(4)
   p(5)
   p(6)
   p(7)
101 f(1)                                    SD=20
            10              920
            20              800
            30              750
            40              630
            50              610
            60              530
            70              520
            90              380
           110              350
           150              200
```

WinSAAM **model** of the study for Subject 1

Fit of our model to **Subject** 1's data in SemiLog and Linear formats exposes a **single phase disposition**



Best estimates of values for the adjustable parameters and their errors .. note fsd=sd/mean. What is the %'ge **error** for each parameter?

```
> fsd(i)
* VALUES MAY NOT RELATE TO CURRENT PARAMETERS
* L ( 0, 1)    1.026E-02     FSD( 1)   3.637E-02
* K ( 1)       1.002E-01     FSD( 2)   1.768E-02
```

```
 cd "demonstrations\Multiple studies 1
c:\stata2019\New Biostatistics Program for Penn Vet aug 12 2020\demonstrations\Multiple studies 1

. dir *.stmp
   0.6k   9/01/20 10:37  cor.stmp
   0.7k   9/01/20 10:37  cov.stmp
   6.2k   9/01/20 10:37  data.stmp
   0.4k   9/01/20 10:37  filenames.stmp
   1.8k   9/01/20 10:37  params.stmp
   3.6k   9/01/20 10:37  partials.stmp
   0.4k   9/01/20 10:37  popcom.stmp
   0.1k   9/01/20 10:37  popcor.stmp
   0.1k   9/01/20 10:37  popcov.stmp
   0.1k   9/01/20 10:37  popparam.stmp

. insheet using params.stmp, clear
(11 vars, 36 obs)

. gen order=_n
. describe

Contains data
  obs:            36
 vars:            12
-------------------------------------------------------------
              storage   display    value
variable name   type    format     label      variable label
-------------------------------------------------------------
label           str7    %9s                   Label
study           byte    %8.0g                 Study
studyname       str9    %9s                   Study Name
parameter       str8    %9s                   Parameter
dest            byte    %8.0g                 Dest
src             byte    %8.0g                 Src
form            str1    %9s                   Form
value           float   %9.0g                 Value
min             byte    %8.0g                 Min
max             int     %8.0g                 Max
fsd             float   %9.0g                 FSD
order           float   %9.0g
-------------------------------------------------------------
Sorted by:
     Note: Dataset has changed since last saved.
```

In course **Module 1** we ran a **Population Investigation**.
Ten **tab**-delimited**, mnemonically**-identified text files were **automatically** created by the **Population** Studies invocation. This presentation will **access** the contents of 4 of these.

It rarely wastes time by creating a '**ledger**' of initial data order in (moderately) complicated datasets. This step assures that we **can** always return to the original sort order of the data

First stmp file opened is the '**params.stmp**' file. As always we start by describing the data … note there are **12 variables and 36 observations** in this dataset … The observation count compiles from 9 parameters (next slide) for each of the 4 subjects

```
. sort study order
. table label study, c(mean value) col format(%12.3f)
```

Label	Study 1	2	3	4	Total
K1	0.100	0.102	0.098	0.050	0.087
L01	0.010	0.021	0.019	0.010	0.015
p1_Vd	9.978	9.835	10.223	19.946	12.496
p2_Cl	0.102	0.202	0.195	0.196	0.174
p3_Cmax	1002.200	1016.700	978.180	501.350	874.607
p4_t1/2	67.584	33.683	36.403	70.620	52.073
p5_AUC	97714.000	49407.000	51372.000	51079.000	62393.000
p6_AUMC	9527300.000	2401000.000	2698000.000	5204100.000	4957600.000
p7_MRT	97.503	48.595	52.518	101.880	75.124

Here we show the **subject** level diversity of the adjustments achieved by the **combined** Marquardt Gauss Newton optimizer, as the best possible fit is found.
Note that we have **extracted** the variables label, study, and value from which we create this initial subject level account

```
. table study param, c(mean value) row format(%9.3f)
```

Study	Parameter K(01)	L(00,01)	P(01)	P(02)	P(03)	P(04)	P(05)	P(06)	P(07)
1	0.100	0.010	9.978	0.102	1002.200	67.584	97714.000	9.53e+06	97.503
2	0.102	0.021	9.835	0.202	1016.700	33.683	49407.000	2.40e+06	48.595
3	0.098	0.019	10.223	0.195	978.180	36.403	51372.000	2.70e+06	52.518
4	0.050	0.010	19.946	0.196	501.350	70.620	51079.000	5.20e+06	101.880
Total	0.087	0.015	12.496	0.174	874.607	52.073	62393.000	4.96e+06	75.124

Another perspective of the same information.
Note that the variable **Parameter** has been substituted for the variable **Label** illustrating flexibility here.

```
. insheet using popparam.stmp, clear
. list

     +-------------------------------------------------+
     | label   parame~r   popmean   esterror     popstd |
     |-------------------------------------------------|
  1. |   L01   L(00,01)   .014894   .0024565   .0048927 |
  2. |    K1      K(01)   .082357    .013374    .022627 |
     +-------------------------------------------------+

(5 vars, 2 obs)

. more
. insheet using data.stmp, clear
(13 vars, 68 obs)

. keep label study c category time qc qo sd fsd
. table label study if c==0, c(mean qc mean fsd) col
-------------------------------------------------------------
          |                      Study
    Label |        1         2         3         4     Total
----------+--------------------------------------------------
    p1_Vd |   9.9784    9.8355    10.223    19.946  12.49572
          |  .014209   .019752  .0090487  .0041916  .0118003
          |
    p2_Cl |   .10234     .2024    .19466    .19577  .1737925
          |  .022016   .015914  .0078634  .0067731  .0131416
          |
  p3_Cmax |   1002.2    1016.7    978.18    501.35  874.6075
          |   .03038     .3836    .18308   .017906  .1537415
          |
  p4_t1/2 |   67.584    33.683    36.403     70.62   52.0725
          |  .036008   .035523     .0168   .010894  .0248062
          |
   p5_AUC |    97714     49407     51372     51079     62393
          |  .012922    .38156    .18217   .016004   .148164
          |
  p6_AUMC |  9527300   2401000   2698000   5204100   4957600
          |  .044646    .38283    .18281   .020685  .1577428
          |
   p7_MRT |   97.503    48.595    52.518    101.88    75.124
          |  .036008   .035523     .0168   .010894  .0248062
-------------------------------------------------------------
```

Loading the **popparam** stmp file gives us access to the **population** estimates of the combined subject level results. Thus was we can compare subject values with the population values and display the results graphically. These statistical values are characteristic of the population.

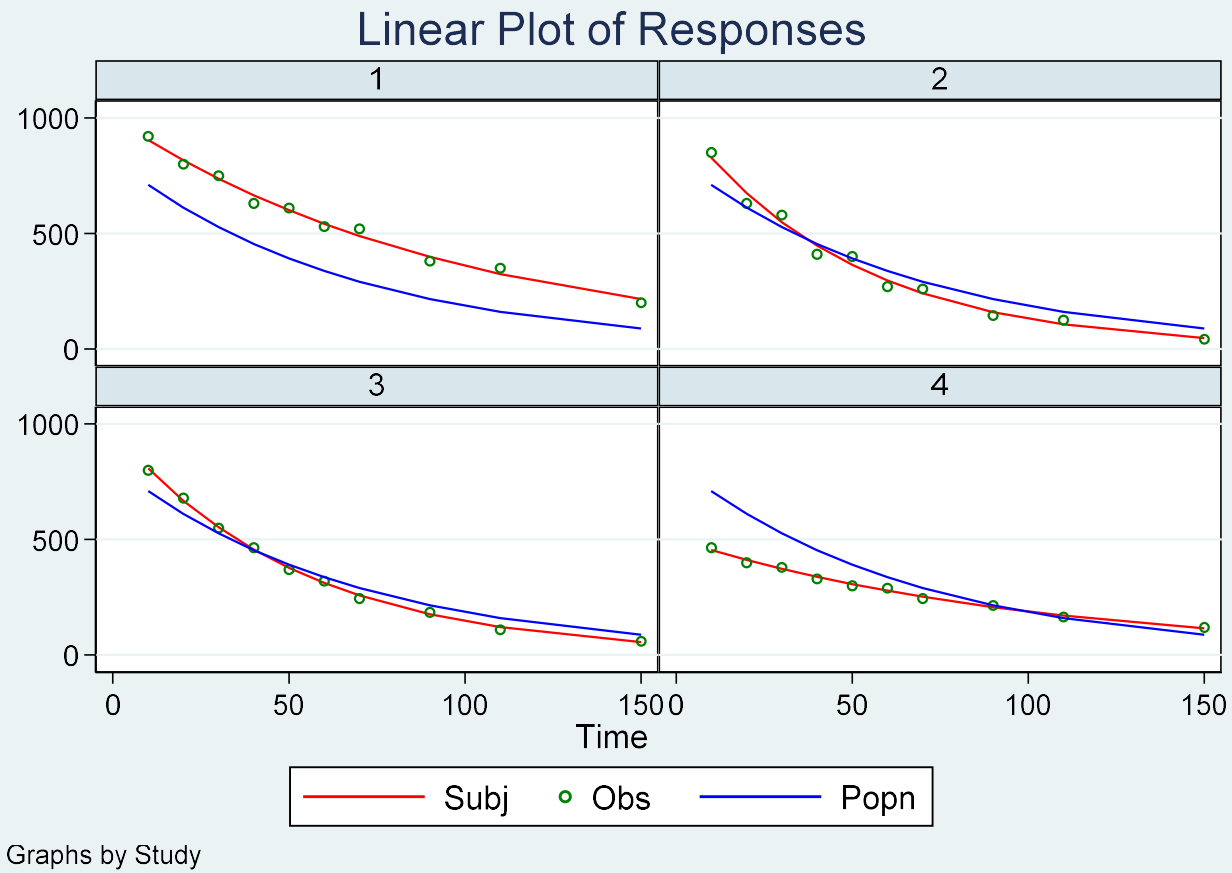
This table (data.stmp) shows the estimates and errors (fsd's) of the **subject** level parameters. Thus enabling us to locate subjects whose responses may be at **greatest** variance compared with the remaining members of the study population.

Note the somewhat large fsd values for Cmax, AUC, and AUMC in study 2. An fsd of **50%** for a subject's estimate of a parameter would suggest a subject for whom the parameter in question is **not identified**

```
. gen ms_vals=10000*0.082357*exp(-0.014894*time) // Takes care of dose

. scatter qc qo ms_vals time if c==1, c(l . l) ms(i oh i) ///
   by(study, style(compact) ti(Linear Plot of Responses)) lc(red . blue) mcolor(. green .)  ///
   ylabel(,angle(0)) legend(row(1) lab(1 "Subj") lab(2 "Obs") lab(3 "Popn"))

. more
```



Here we present a linear trellis plot of

1) Subject observations(Green dots),
2) The fit of the response to these observations (Red line).
3) The population mean response (Blue line)

Note that the line

gen ms_vals=10000*0.082357*exp(-0.014894*time)

serves to generate population predictions by manually extracting K1 and L0,1 from the population table. Two slides back

```
. scatter qc qo ms_vals time if c==1, c(l . l) ms(i oh i) yscale(log) ///
   by(study, style(compact) ti(Semi Log Plot of Responses)) lc(red . blue) mcolor(. green .) ///
   ylabel(100 300 1000,angle(0)) legend(row(1) lab(1 "Subj") lab(2 "Obs") lab(3 "Popn"))

. more
```

Here we produce another trellis plot of the disposition of the drug from each subject. This time using the plot option ,**yscale**, to set the yscale to **logarithmic** and a reasonable ylabeling scheme this trellis is presented in semilog fashion.

Note the **reasonable** agreement between the population mean response and the subject specific response for subjects 2,3, and 4 but for subject 4 the subject and population plots are significantly in disagreement.

```
. insheet using popcom.stmp, clear
(6 vars, 8 obs)

. sort label study

. list , sepby(label)
```

	label	parame~r	study	studyname	studyv~e	popmean
1.	K1	K(01)	1	SUBJECT_1	.10022	.082357
2.	K1	K(01)	2	SUBJECT_2	.10167	.082357
3.	K1	K(01)	3	SUBJECT_3	.097818	.082357
4.	K1	K(01)	4	SUBJECT_4	.050135	.082357
5.	L01	L(00,01)	1	SUBJECT_1	.010256	.014894
6.	L01	L(00,01)	2	SUBJECT_2	.020578	.014894
7.	L01	L(00,01)	3	SUBJECT_3	.019041	.014894
8.	L01	L(00,01)	4	SUBJECT_4	.0098151	.014894

This small table (**popcom**) shows how well the subject level parameters agree, or disagree with the population estimates. **K(01)** for subject 2 is considerably higher that the population mean, and, L(0,1) is quite different from the mean for the same subject

Stata Code to Capture Multiple Studies Demonstration

Setting working directories

```
cd "c:\stata2019\New Biostatistics Program for Penn Vet aug 12 2020"
cd "demonstrations\Multiple studies 1
dir *.stmp
```

Parameters file
Leads to tables

```
insheet using params.stmp, clear
gen order=_n
describe
sort study order
list in 1/9 , sepby(study)
list in -9/l
table label study, c(mean value) col format(%12.3f)
table study param, c(mean value) row format(%9.3f)
more
insheet using popparam.stmp, clear
list
more
```

Data file
Leads to plots

```
insheet using data.stmp, clear
keep label study c category time qc qo sd fsd
table label study if c==0, c(mean qc mean fsd) col
gen ms_vals=10000*0.082357*exp(-0.014894*time) // Takes care of dose
scatter qc qo ms_vals time if c==1, c(l . l) ms(i oh i) ///
  by(study, style(compact) ti(Linear Plot of Responses)) lc(red . blue) mcolor(. green .) ///
  ylabel(,angle(0)) legend(row(1) lab(1 "Subj") lab(2 "Obs") lab(3 "Popn"))
more
scatter qc qo ms_vals time if c==1, c(l . l) ms(i oh i) yscale(log) ///
  by(study, style(compact) ti(Semi Log Plot of Responses)) lc(red . blue) mcolor(. green .) ///
  ylabel(100 300 1000,angle(0)) legend(row(1) lab(1 "Subj") lab(2 "Obs") lab(3 "Popn"))
more
```

Subject comparisons
Listed

```
insheet using popcom.stmp, clear
sort label study
list , sepby(label)
```

Edited state of our working directory (folder) in conjunction with a population study invocation

```
. dir
  <dir>   9/01/20 15:04  .
  <dir>   9/01/20 15:04  ..
   4.2k   9/01/20 10:37  assemble.txt
---------------------------------------------------------
   0.5k   9/01/20 10:41  PK_GW4_Pop_1.sprj
---------------------------------------------------------
   1.0k   9/01/20 10:28  GW4-1_SUB_1.saam
   1.0k   9/01/20 10:28  GW4_1_SUB_2.saam
   1.0k   9/01/20 10:34  GW4_1_SUB_3.saam
   1.0k   9/01/20 10:36  GW4_1_SUB_4.saam
---------------------------------------------------------
   0.6k   9/01/20 10:37  cor.stmp
   0.7k   9/01/20 10:37  cov.stmp
   6.2k   9/01/20 10:37  data.stmp
   0.4k   9/01/20 10:37  filenames.stmp
   0.4k   9/01/20 10:37  popcom.stmp
   0.1k   9/01/20 10:37  popcor.stmp
   0.1k   9/01/20 10:37  popcov.stmp
   0.1k   9/01/20 10:37  popparam.stmp
   1.8k   9/01/20 10:37  params.stmp
   3.6k   9/01/20 10:37  partials.stmp
---------------------------------------------------------
   1.1k   9/01/20 11:34  Display Results MS 1.do
  65.9k   9/01/20 15:04  Multiple Studies 1.pptx
```

Multiple Studies Text Output

Project control file

Project files

Population study files (stmp)

Stata do file and This PPT file

A Quick Look at Automation … the Critical Tool for the Analysis of LARGE STUDIES

Interoperability as WinSAAM and Stata synergize

Ray Boston, 2004

Steps In Automation

```
cd "C:\Stata2019\New Biostatistics Program for Penn Vet aug 12 2020\"
cd "demonstrations\Multiple Studies 1
use "plasma conc dec 5 2020", clear
egen g=group(id)
qui su g
local lmax=r(max)
tempname fh fw
```

Preamble

```
forval i=1/`lmax' {
preserve
file open `fh' using "GW4_P490_Ref.saam", read
keep if g==`i'
sort time
local name="NT_"+string(id[1])
cap conf file "`name'.saam"
if _rc==0 {
  erase "`name'.saam"
  }
```

Setting up Reference model

```
file open `fw' using "`name'.saam", write text
local line1="A SAAM31"
file write `fw' "`line1'"
file write `fw' _col(42)
file write `fw' "`name'"
file write `fw' _n
file read  `fh' line
while r(eof)==0 {
  di " Line: `line'"
  file read `fh'  line
  file write `fw' "`line'"
  file write `fw' _n
}
```

Adding the reference model to the subject file: id: 1, 2, 3, and 4

```
format time %4.2f
format obs  %6.2f
file write `fw' "101" _col(42) "fsd=.08" _n
local N=_N
forval j=1/`N' {
  local tt=time[`j']
  local pp=obs[`j']
  file write `fw' _col(13) (`tt') _col(27) (`pp') _n
}
```

Adding subject data to subject file

```
sort time
file close `fw'
file close `fh'
restore
}
```

Closing files

Reference Model

```
2       10
H PAR
C
c# K1=K(1)
c# L01=L(0,1)
c# p1_Vd=p(1)
c# p2_Cl=p(2)
c# p3_Cmax=p(3)
c# p4_t1/2=p(4)
c# p5_AUC=p(5)
c# p6_AUMC=p(6)
c# p7_MRT=p(7)
C
   K(1)      1.002162E-01  0.000000E+00   1.000000E+04
   L(0,1)    1.025612E-02  0.000000E+00   1.000000E+04
   ic(1)    10000
 p(1)=1/k(1)
 p(2)=p(1)*l(0,1)
 p(3)=ic(1)*k(1)
 p(4)=alog(2)/l(0,1)
 p(5)=k(1)*ic(1)/l(0,1)
 p(6)=k(1)*ic(1)/l(0,1)/l(0,1)
 p(7)=p(6)/p(5)
H DAT
   p(1)
   p(2)
   p(3)
   p(4)
   p(5)
   p(6)
   p(7)
```

Data for subject 1

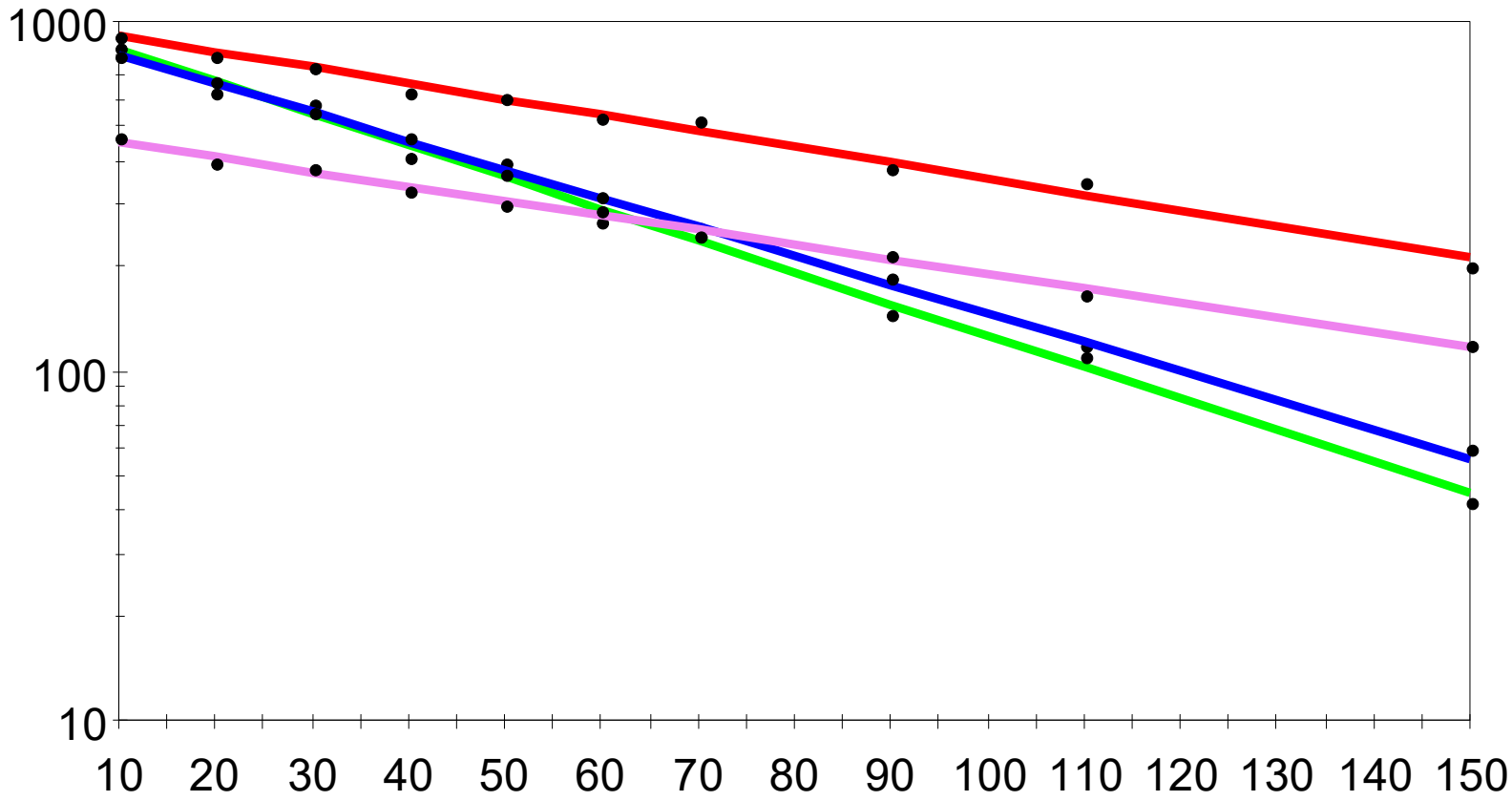
	time	obs	id
1.	10	920	1
2.	20	800	1
3.	30	750	1
4.	40	630	1
5.	50	610	1
6.	60	530	1
7.	70	520	1
8.	90	380	1
9.	110	350	1
10.	150	200	1

```
A SAAM31                                        NT_1
H PAR
C
c# K1=K(1)
c# L01=L(0,1)
c# p1_Vd=p(1)
c# p2_Cl=p(2)
c# p3_Cmax=p(3)
c# p4_t1/2=p(4)
c# p5_AUC=p(5)
c# p6_AUMC=p(6)
c# p7_MRT=p(7)
C
   K(1)      1.011125E-01  0.000000E+00   1.000000E+04
   L(0,1)    1.048873E-02  0.000000E+00   1.000000E+04
   ic(1)    10000
 p(1)=1/k(1)
 p(2)=p(1)*l(0,1)
 p(3)=ic(1)*k(1)
 p(4)=alog(2)/l(0,1)
 p(5)=k(1)*ic(1)/l(0,1)
 p(6)=k(1)*ic(1)/l(0,1)/l(0,1)
 p(7)=p(6)/p(5)
H DAT
   p(1)
   p(2)
   p(3)
   p(4)
   p(5)
   p(6)
   p(7)
101                                             fsd=.08
            10            920
            20            800
            30            750
            40            630
            50            610
            60            530
            70            520
            90            380
           110            350
           150            200
```

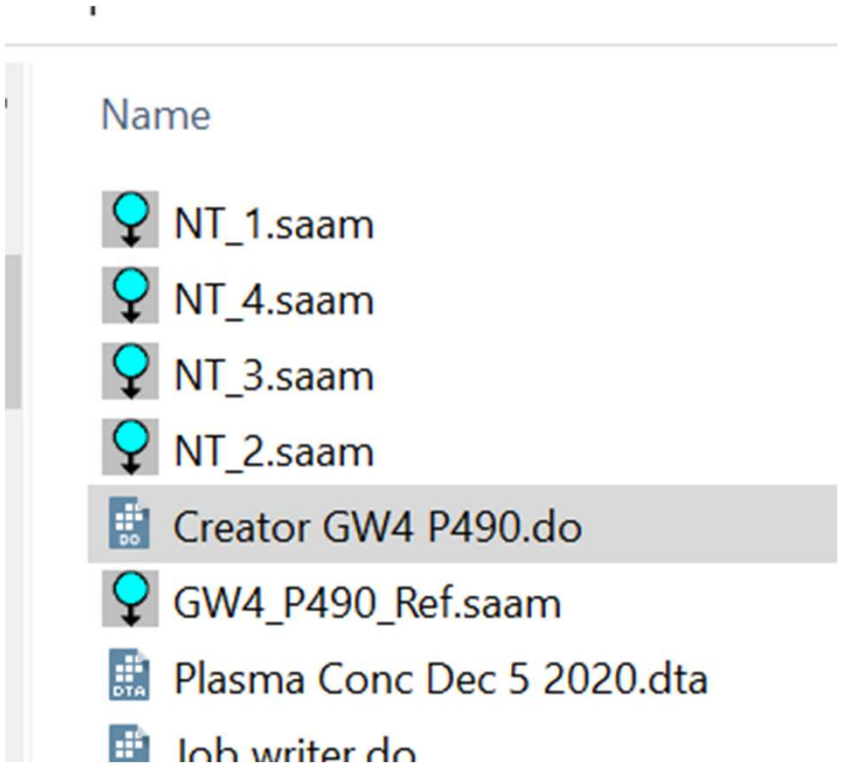
Built WinSAAM ‘decks’ for Subjects 1 … 4

```
A SAAM31                                        NT_4
H PAR
C
c# K1=K(1)
c# L01=L(0,1)
c# p1_Vd=p(1)
c# p2_Cl=p(2)
c# p3_Cmax=p(3)
c# p4_t1/2=p(4)
c# p5_AUC=p(5)
c# p6_AUMC=p(6)
c# p7_MRT=p(7)
C
   K(1)      4.960401E-02  0.000000E+00   1.000000E+04
   L(0,1)    9.632042E-03  0.000000E+00   1.000000E+04
   ic(1)    10000
 p(1)=1/k(1)
 p(2)=p(1)*l(0,1)
 p(3)=ic(1)*k(1)
 p(4)=alog(2)/l(0,1)
 p(5)=k(1)*ic(1)/l(0,1)
 p(6)=k(1)*ic(1)/l(0,1)/l(0,1)
 p(7)=p(6)/p(5)
H DAT
   p(1)
   p(2)
   p(3)
   p(4)
   p(5)
   p(6)
   p(7)
101                                             fsd=.08
            10            465
            20            400
            30            380
            40            330
            50            300
            60            290
            70            245
            90            215
           110            165
           150            120
```

Model fits to each subject (red:1, green: 2, blue:3, violet: 4)



Critical Elements Linked to Automation



Stata command file
WinSAAM reference model
Data file

Thank You