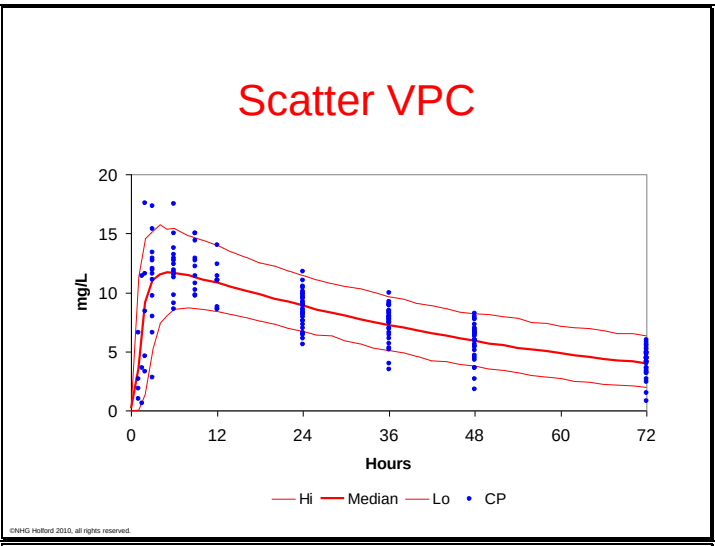
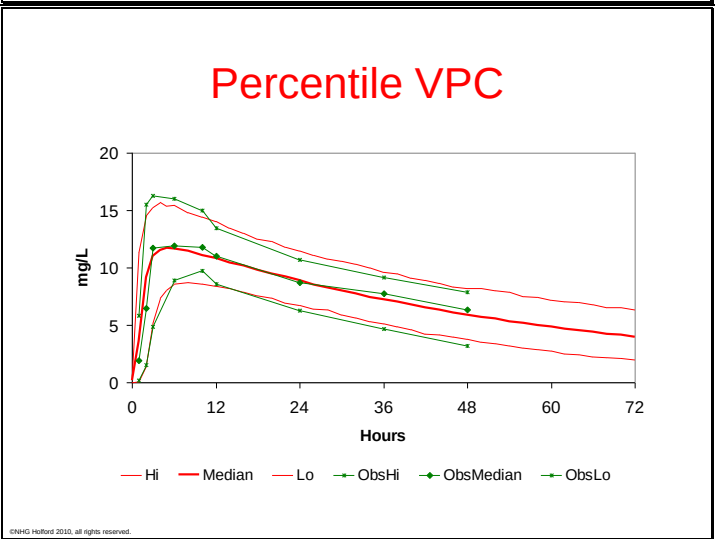
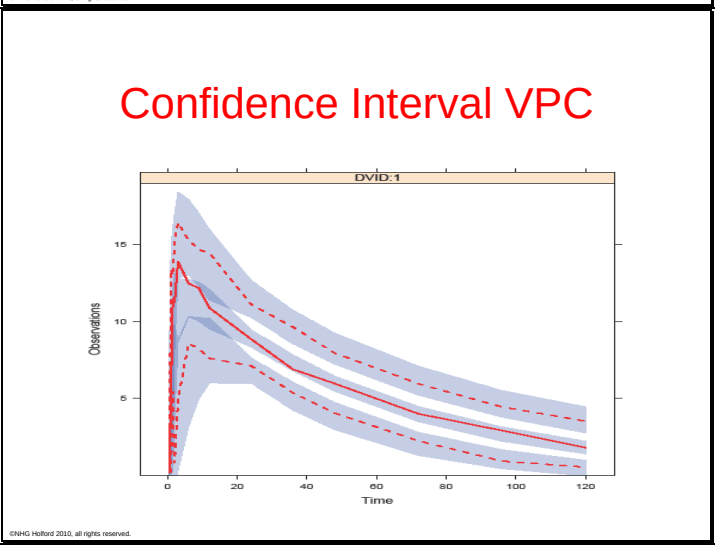
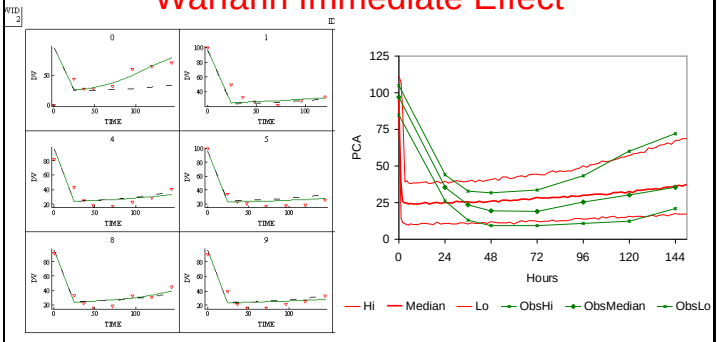
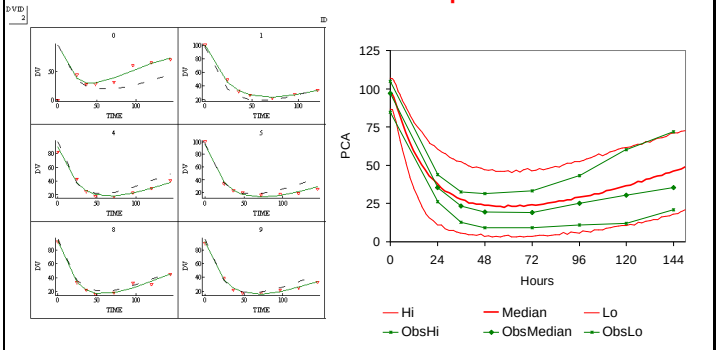
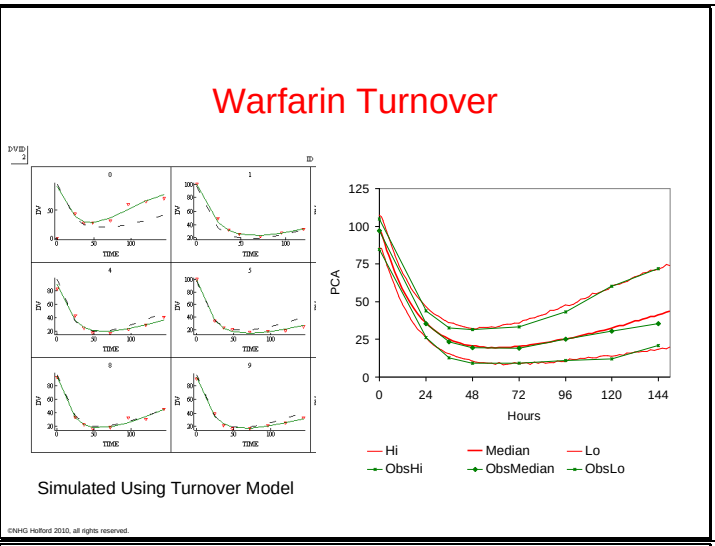
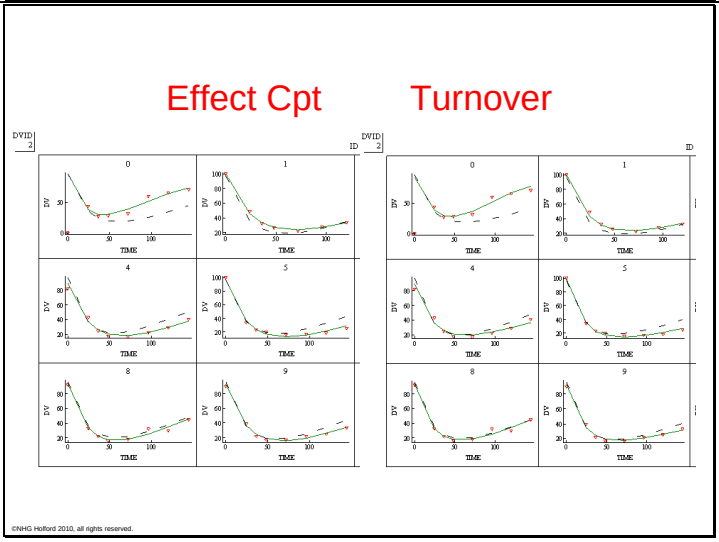
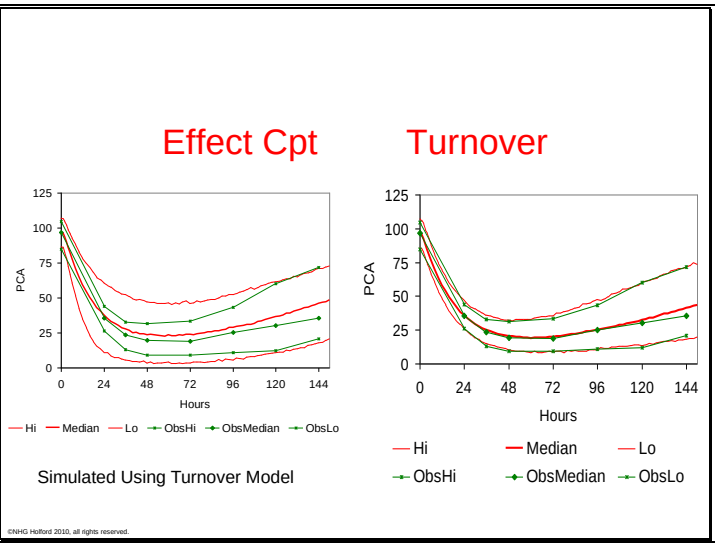
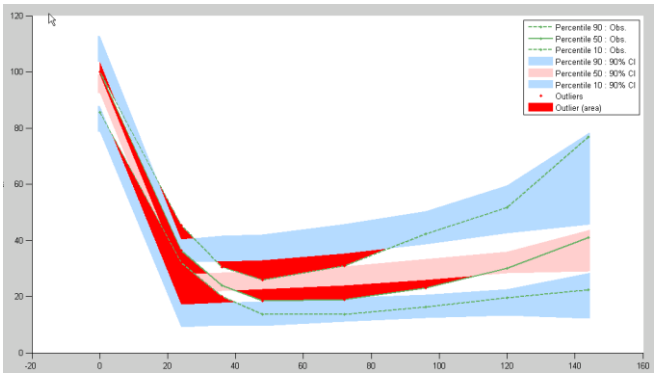
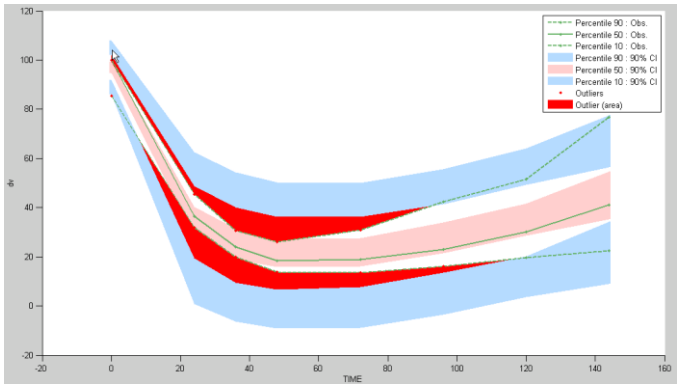
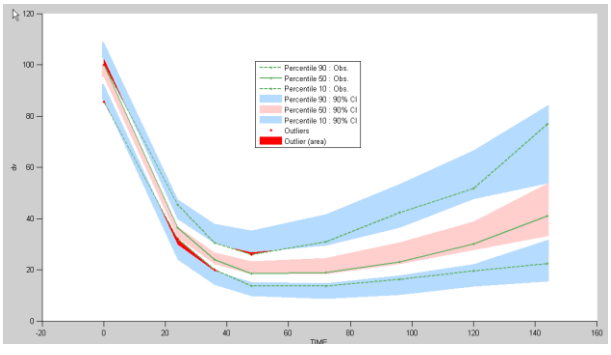
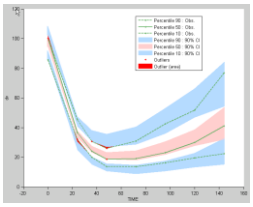
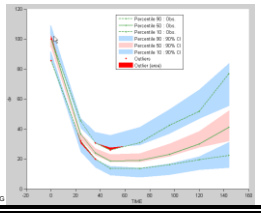
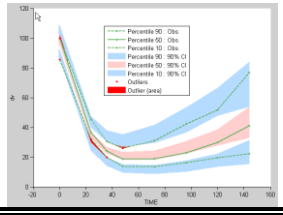
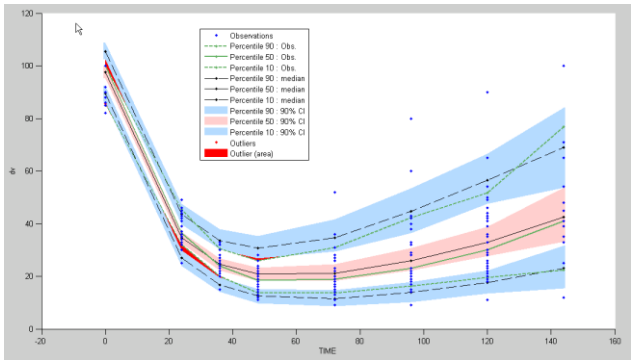
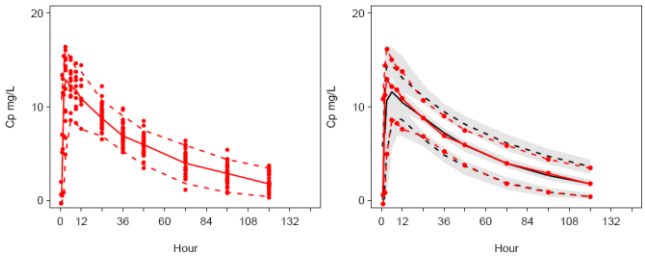
Slide 1	# An Introduction to # Visual Predictive Checks Nick Holford University of Auckland	
Slide 2	# Outline • What is a Visual Predictive Check? • What choices are there in presentation? • What can it help to show? • What may it fail to show? 	
Slide 3	# What is a VPC? • Comparison of Observations and Simulated Predictions (SPRED) – Simulated predictions include fixed and random between subject variability as well as residual error – They are different from 'population' predictions (PRED) (fixed effects without random effects) and individual predictions (IPRED)(shrinkage) • VPC compares statistics derived from the distribution of observations and the distribution of SPRED at specific times or time intervals – E.g. median and 90% intervals at 1 h after the dose – Intervals can be joined together in time sequence to create bands (but most often the bands are called 'intervals') 	VPCs use a different kind of prediction compared with traditional diagnostic plots. They are based on simulations of model predictions including random effects (especially between subject variability (BSV)). Summary measures of the distribution of predictions and observations are compared visually. Typical summary measures are the median and an interval defined by the lower 5% and upper 5% of the values.

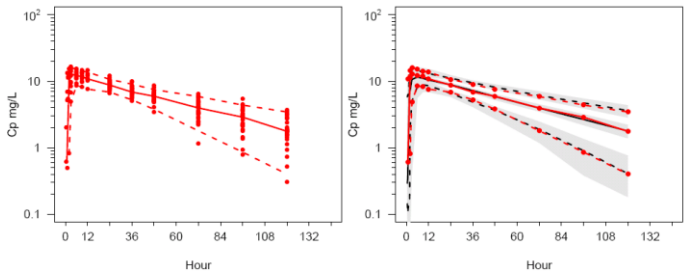
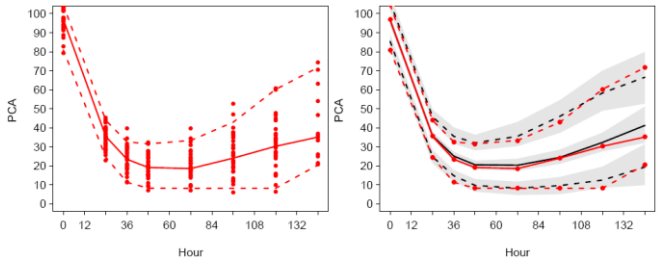
Slide 4		As you have seen there are many ways of creating VPCs with increasing complexity. In summary there are three basic kinds of VPC. The first is the scatter plot VPC which shows the observations along with some simple prediction intervals. This is a useful starting point for connecting observations with prediction intervals. However when there is a lot of data the actual distribution of the observations can be hard to appreciate.
Slide 5		The second kind of VPC summarises the distribution of observations with observation intervals so they can be compared directly with the prediction intervals and the medians of the observed and predicted values. The percentile VPC is easier to interpret when there are lots of observations.
Slide 6		The third type shows the 95% confidence interval around each of the prediction intervals obtained by simulation. The red lines are the observation intervals.

Slide 7	# What Can a VPC Show? 	
Slide 8	## Warfarin Immediate Effect  Simulated Using Turnover Model 	Data has been simulated from a warfarin PKPD model involving turnover of prothrombin complex activity (PCA) after a single oral dose of warfarin. The PK model is first order absorption and first order elimination from one compartment. The data has been fitted with the same PK model used to simulate the data but the PD model assumes an immediate effect of warfarin plasma concentration on PCA. The left hand plots show individual predictions obtained from empirical Bayes estimates and the corresponding observations. The right hand plot is a percentile VPC with 90% intervals. Both the individual plots and the VPC show poor predictions.
Slide 9	## Warfarin Effect Compartment  Simulated Using Turnover Model 	The PD model now assumes a delayed onset of warfarin effect using an effect compartment model for concentrations driving the change in PCA. The individual predictions look very good but the VPC median prediction lies above the median observation from 24 h onwards and the 90% interval is clearly much wider than the observations. This suggests the model is not properly describing the data despite the very good individual predictions.

Slide 10		Finally we can see what happens when the true model is used to fit the data. When a turnover model is used the individual predictions remain good and the VPC percentile plot looks good as well.
Slide 11		Notice that plots using empirical Bayes estimates for predictions are essentially the same for both the effect compartment and turnover model. Yet the VPCs show the predictions of the effect compartment model are a poorer description of the observations. This is a consequence of shrinkage. The shrinkage for the effect cpt model was 14-44% and for the turnover model was 11-40%. It has been suggested that all parameters must have less than 20% shrinkage in order to draw reliable conclusions from individual plots. However, shrinkage estimates do not take account of the correlation between parameters and it is the correlated set of parameters that determines the prediction. Even if one or more parameters have relatively high shrinkage the individual prediction may be quite reliable e.g. for sequential PKPD models using the IPP approach.
Slide 12		This slide compares the VPC using the incorrect effect compartment model with the VPC obtained from the true turnover model. It is reassuring to see that the VPC with the true model has good agreement with the observations.

Slide 13	Monolix 3.2 Immediate  		
Slide 14	Monolix 3.2 Effect Compartment  		
Slide 15	Monolix 3.2 Turnover PKPDLIB  		

Slide 16	**Monolix 3.2 Turnover** MLX-TRAN CMT  CMT =S0,Imax,C50,kout CMTlib=Rin,Imax,C50,kout MLX-TRAN CMTlib  PKPD Library 	
Slide 17	**Monolix 3.2 Turnover**  ©NHG Holford 2010, all rights reserved.	
Slide 18	**nmvpc** **KA1L**  ©NHG Holford 2010, all rights reserved.	

Slide 19	nmvpc KA1L  		
Slide 20	nmvpc Turnover  		
Slide 21	How to do a VPC • Simulate Data – Can be the hardest part – Simulation times (binning)? – How to simulate covariates? • Group Simulated Predictions at each time – Needs some programming • Group Observations at each time – Needs some programming 		There are 3 steps involved in creating a VPC. The first step is to simulate from the model to produce predictions. This step typically requires user intervention for every dataset that is being studied. The next two steps can usually be automated with procedures that are the same for all problems. A convention is needed to identify the independent and dependent variables (especially when there is more than one type of observation e.g. Concentrations and efffects).

Slide 22

nmvpc

- WFN command
  - nmvpc.bat
- R scripts
  - nmvpc_PKPD.R
  - nmvpc_PKPD_Functions.R
- R scripts can be run directly using R
- nmvpc helps to automate repetitive features of doing VPC e.g. with covariate selection

Slide 23

NM-TRAN

```
$PK
; simulation start
OBS=DV
; simulation end

[Usual $PK (or $PRED) code]

;Simulation Start
;must produce NONMEM table file with REP,ID,TIME,DVID,DV,MDV,PRED
REP=IREP

$SIM (20091125) ONLYSIM NSUB=10
$TABLE REP ID TIME DVID DV MDV PRED OBS SEX ; SDY TAD DOSE
NOAPPEND ONEHEADER NOPRINT FILE=vpc.fit
;Simulation End
```


Slide 24

VPC Using WFN and R

- WFN modification
  - Include path to R.exe in wfn.bat

```
:orgpath
rem ******* Set up Paths to Other Software ******
rem If you use R then set RPATH
set RPATH=C:\Apps\R-2.10.1\bin
rem ******* End Check this ********
```

Slide 25

nmvpc_PKPD.R - 1

```
### The following items may be modifed by runVPC.bat
modelName   = "ka1_to_emax1_simln"       # No file extension
obsFileName = "ka1_to_emax1_simln.csv"   # Observation data
percentile  = 0.9                        # Percentile for prediction intervals

logaxis = "y" # choose axis for log scale (use "x", "y", "xy" or "")

Xlabel  = "Hour" # X-axis label on graph
Xmin    =   0 # minimum scale for x-axis
Xmax    = 144 # maximum scale for x-axis
Xtick   =  12 # ticks on x-axis at these intervals for linear scale

thisDVID=1 # #identify observation type using the DVID variable in the
simulation and observation data files

Ylabel  = "Cp mg/L" # Y-axis label on graph
Ymin    =  0 # minimum scale for y-axis
Ymax    = 50 # maximum scale for y-axis
Ytick   =  5 # ticks on y-axis at these intervals for linear scale

pdfTxt  = "" # pdf file name identifier e.g. use with covariate selection
```


Slide 26

nmvpc_PKPD.R - 2

```
###  The following items may only be changed here in the R script
ISWFN     = T # NONMEM simulation table file: T= Look in Wings for NONMEM run directory;
F = Look in current directory
NMTBL     = ".fit" # NONMEM table file extension
if (ISWFN) NMDIR=".std" # WFN run directory extension

binsim    = T # T if simulation times are not the same for every subject (Otherwise use
times in simulation file 'as is')
hasMDV    = T # T if observation and simulation data file have MDV data item (otherwise
all records are valid DV)
hasDVID   = T # T if observation and simulation data file have DVID data item (otherwise
no selection of DVID)
hasATIM   = F # T if simulation data file has an actual observation time item in the
simulation file (otherwise ATIM=TIME)
plotCI    = T # T if plot confidence intervals
plotPI    = T # T if plot prediction intervals
addLegend = T # T if add a legend to the plot
omitNeg   = T # T if omit negative simulated values

#pdf output options
figOutputDir = "./" # Directory for pdf file output
isPDF        = T # T if generate PDF output otherwise display plots on screen
#pCols       = c(obscol='gray50',simcol='black',pici='gray30')
pCols        = c(obscol='red',simcol='black',pici='gray90') # 0 is black; 100 is white
```


Slide 27

nmvpc_PKPD.R - 3

```
#pdf output options
figOutputDir = "./" # Directory for pdf file output
isPDF        = T # T if generate PDF output otherwise display plots on screen

pCols        = c(obscol='red',simcol='black',pici='gray90') # 0 is black; 100 is white

# bin simulated values by time intervals based on nominal binning times (binTimes)
if (binsim) {
   occ1=c(0.5,seq(1,8,1)) #
   occ2=c(seq(12,144,12)) #
   binTimes <- c(occ1,occ2)
}
timeScale =1 # 0 means use special variable instead of time
#see below File$TIME =*File$variable
# >0 means use this to scale TIME variable (e.g. 52 to scale years to weeks)
```

Slide 28

nmvpc_PKPD.R - 4

```
#Do not change the next line!
#SELECT specific observations e.g. study number
obsFile = obsFile[obsFile$SEX==1,]
simFile = simFile[simFile$SEX==1,]
```

The covariate selection process can be automated with nmvpc.bat

Slide 29

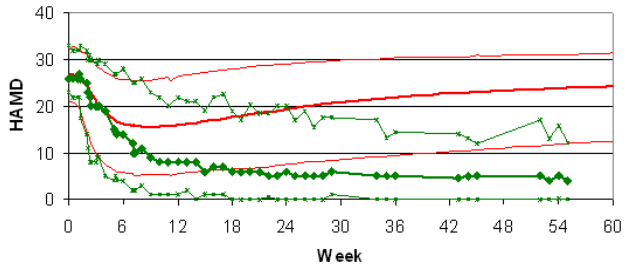
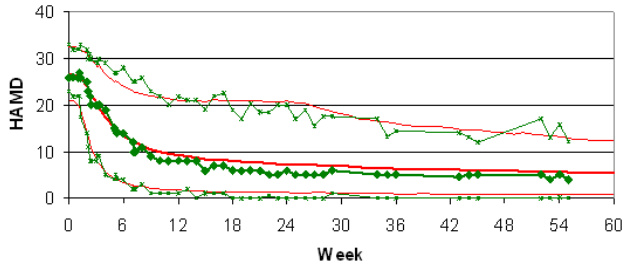
nmvpc.bat - 1

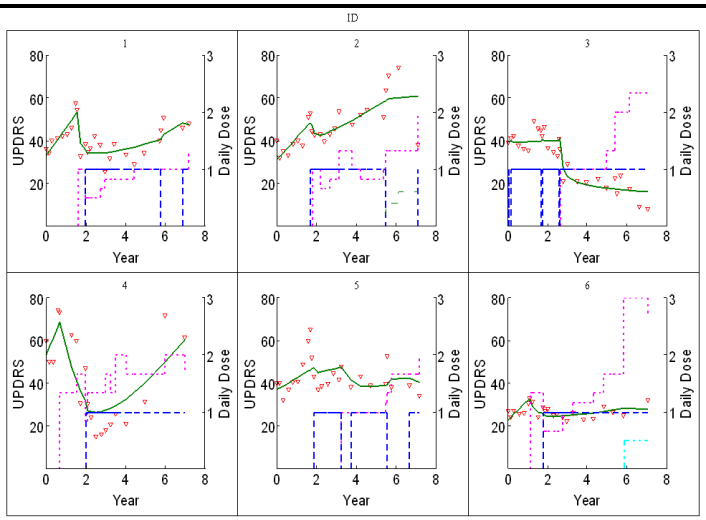
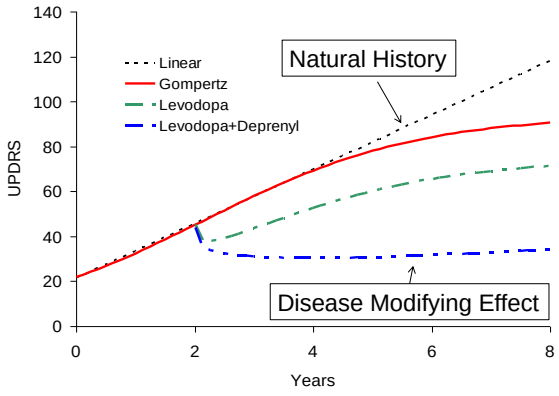
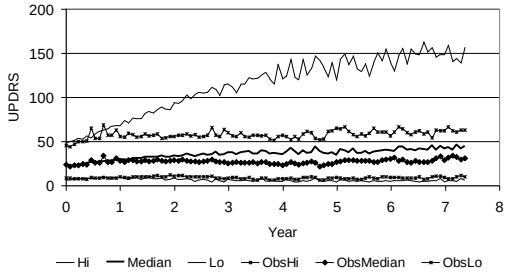
```
rem Set list of models to be simulated
set models=ka1_to_emax1_simln.ctl
set obsfile=ka1_to_emax1_simln.csv
set percentile=0.9

rem Set list VPCs to be created
set vpcnames=CP PCA

goto models

:vpc
set dvidname=%2
set bintimes=0,1,2,3,4,5,6,7,8,12,24,36,48,60,72,84,96,108,120,132,144
goto %2
```

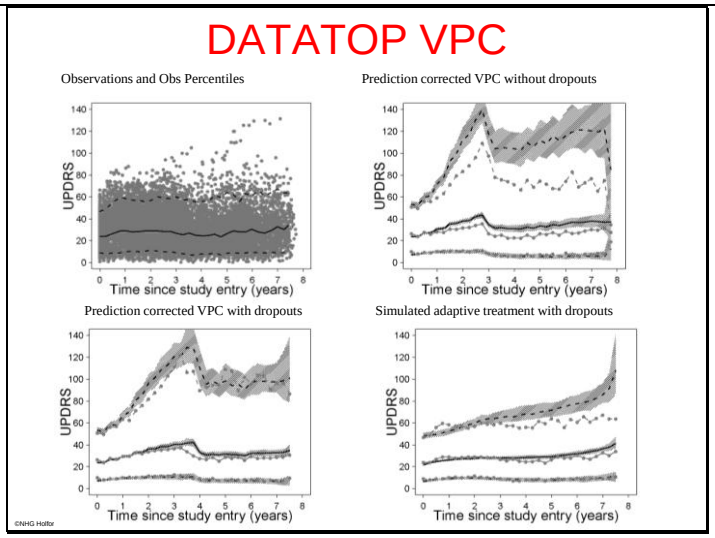

Slide 30

nmvpc.bat - 2

```
rem Define R script variables for each observation type
rem No spaces are allowed for labels.
rem Use '#' which will be replaced by a blank in xlabel and ylabel values

:CP
set logaxis=y
set xlabel=Hour
set xmin=0
set xmax=144
set xtick=12
set dvid=1
set ylabel=Cp#mg/L
set ymin=0
set ymax=100
set ytick=10
set select=
set pdftxt=
goto makeR
```

Slide 31	**VPC – Something Missing?**  Total Dropouts: **23.4%** — Random: 14% — Adverse: 2.0% — Lack of Effect: 6.8% — Remission: 0.6% 	Here is a VPC from a large study of patients in an anti-depressant drug trial. The predicted median and 90% PIs do not agree well with the observed values. This is because there are patients who drop out and the pattern of dropout is influenced by the treatment and patient response. This is known as informative missingness.
Slide 32	**VPC with Dropout**  $h(t)_k = \beta_{0,k} \cdot e^{\beta_{LNT,k} \cdot \ln(time) + \beta_{CE,k} \cdot Ce + \beta_{HAMD,k} \cdot HAMD(time)}$ K=1 (Random), 2 (Adverse), 3 (Lack of Effect), 4 (Remission) Cox E, Veyrat-Follet C, Beal S, Fuseau E, Kenkare S, Sheiner L. A population pharmacokinetic-pharmacodynamic analysis of repeated measures time-to-event pharmacodynamic responses: the antiemetic effect of ondansetron. J Pharmacokin Biopharm 1999;27(6):625-44. 	When it is possible to predict missing values from the data e.g. patients whose HAMD score remains high may dropout because they are not getting better, then this is known as missing at random. A model for the dropout process can be constructed by combining the model predictions for HAMD with a time to event analysis. When the VPC is performed using the dropout model to include a realistic pattern of dropout then the VPC predictions match more closely with the observations. The previous VPC is probably makng good predictions if patients continue with treatment. In that case the observations are 'wrong' because they has been censored by dropouts.
Slide 33	**Parkinson Study Group DATATOP Cohort** Deprenyl and Tocopherol Antioxidative Therapy of Parkinsonism PKPD of anti-parkinsonian treatment and Parkinson's disease over 7 years in 800 patients The Parkinson Study Group. Effect of deprenyl on the progression of disability in early Parkinson's disease. The New England Journal of Medicine 1989;321:1364-1371 	The DATATOP study was performed over 2 year period but patients enrolled in the study were subsequently followed up for 8 years. The time course of disease status in Parkinson's disease and the effects of treatment were described by a disease progress model. The NM-TRAN code for this analysis can be found in Holford et al. 2006. Holford NHG, Chan PL, Nutt JG, Kieburtz K, Shoulson I. Disease progression and pharmacodynamics in Parkinson disease - evidence for functional protection with levodopa and other treatments. J Pharmacokinet Pharmacodyn. 2006 Jun;33(3):281-311.

Slide 34		Disease status was followed with the Unified Parkinson's Disease Response Scale (UPDRS). The UPDRS patterns were quite variable from patient to patient. A major source of variability was the response to individual drug treatments.
Slide 35	Combined Effects of Levodopa and Deprenyl 	The effects of levodopa and deprenyl are shown. Both have offset effects and protective effects which was described by an action on the time constant of a Gompertz asymptotic model. See Holford et al 2006 for details of the model code.
Slide 36	DATATOP VPC  Holford, N. H. G., P. L. Chan, et al. (2006). "Disease progression and pharmacodynamics in Parkinson disease - evidence for functional protection with levodopa and other treatments." J Pharmacokinet Pharmacodyn 33(3): 281-311.	

Slide 37			**Figure 2:** Visual Predictive Check showing (A) observations and observed percentiles, (B) prediction corrected VPC without dropouts and (C) with dropouts, and (D) simulated adaptive treatment with dropouts, Median (solid lines) and 90% prediction intervals (dash-lines) of the observed (joined solid circle lines) and simulated data are shown. The 90% confidence intervals for the median and the prediction intervals are shown in plot as shaded bands.