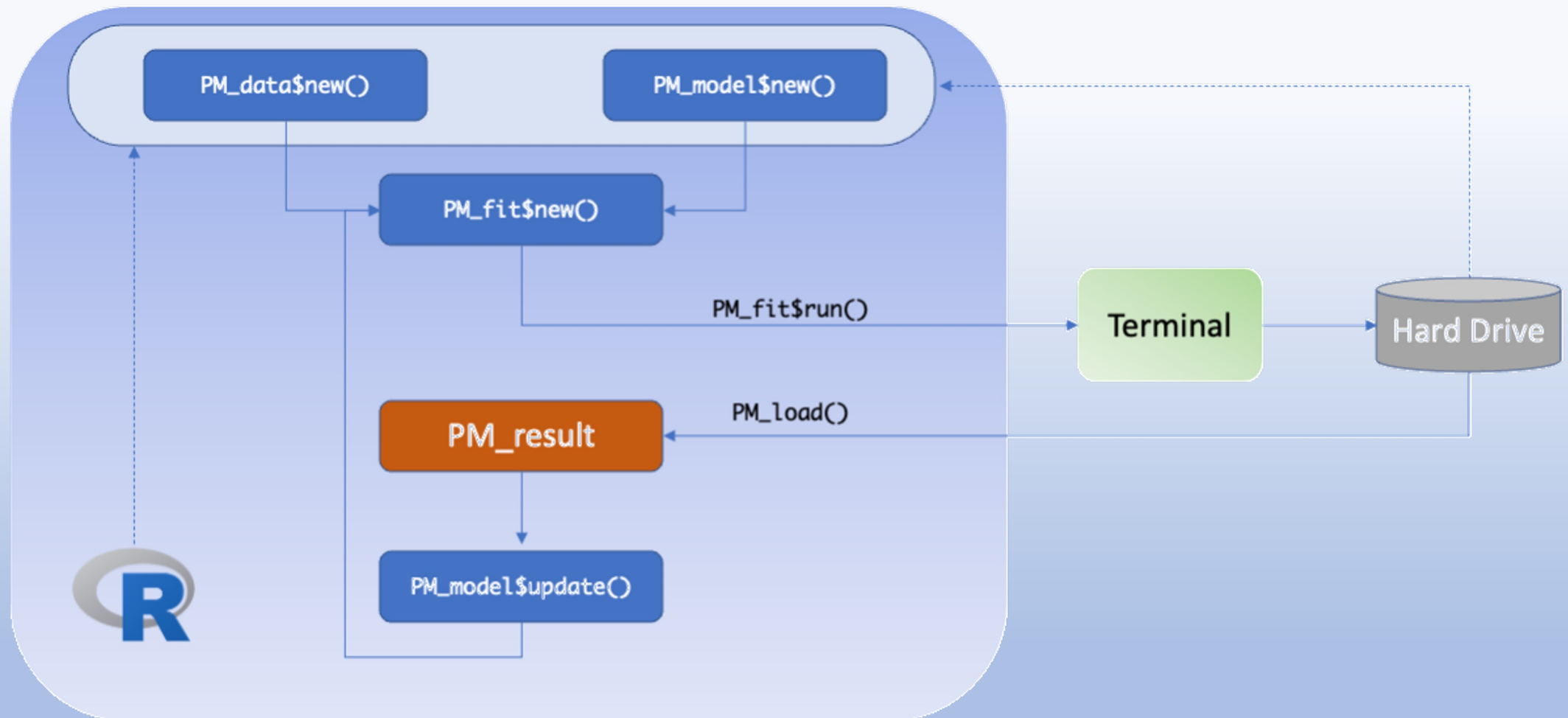# Fitting Data to Models

Michael Neely, MD
Professor of Pediatrics and Clinical Scholar
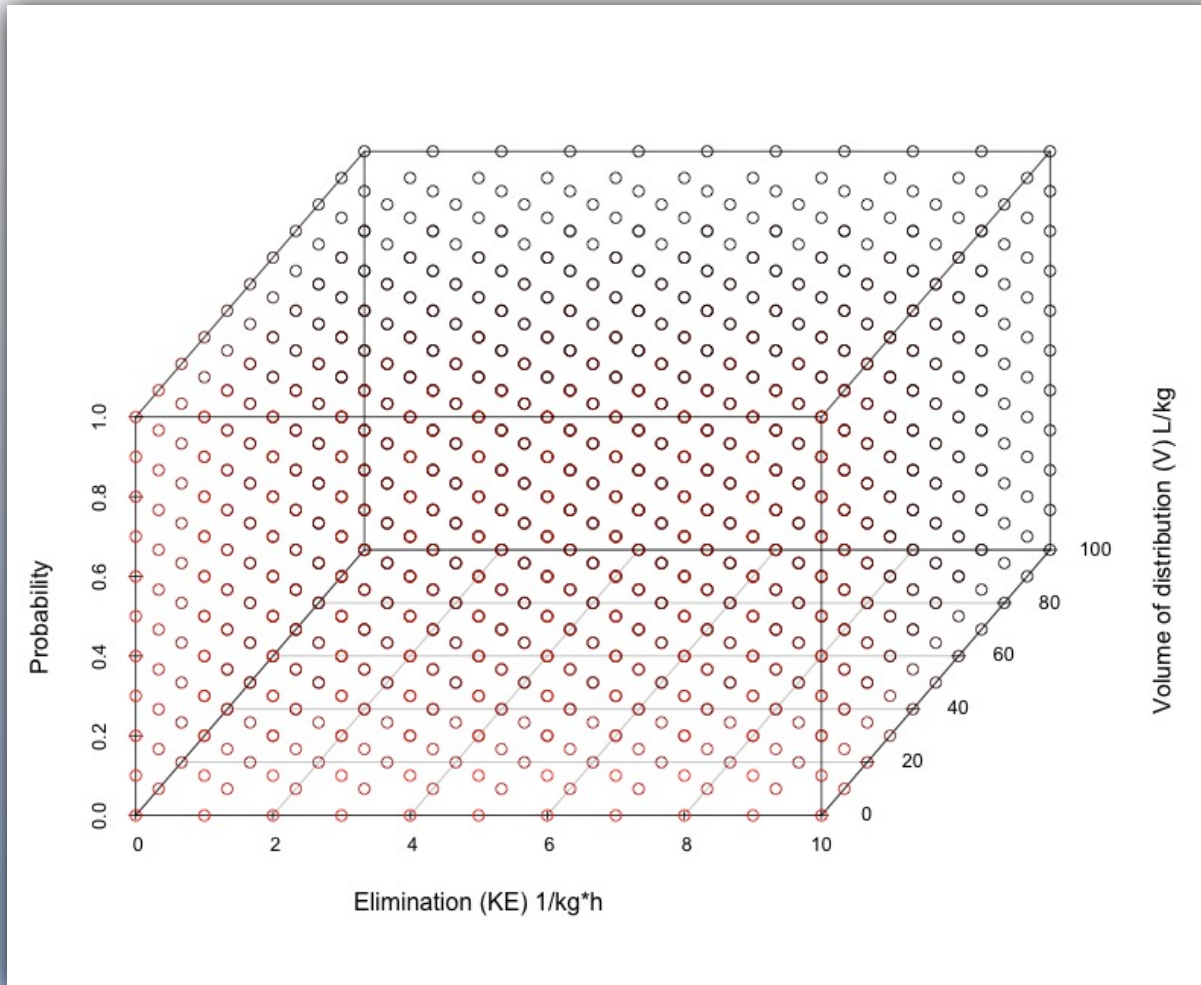Director, Laboratory of Applied Pharmacokinetics and Bioinformatics
University of Southern California, Children's Hospital Los Angeles
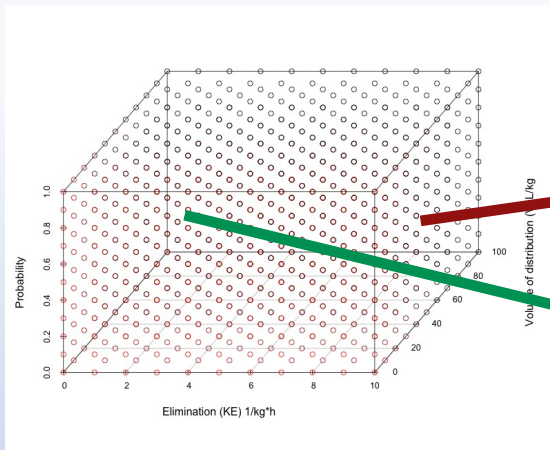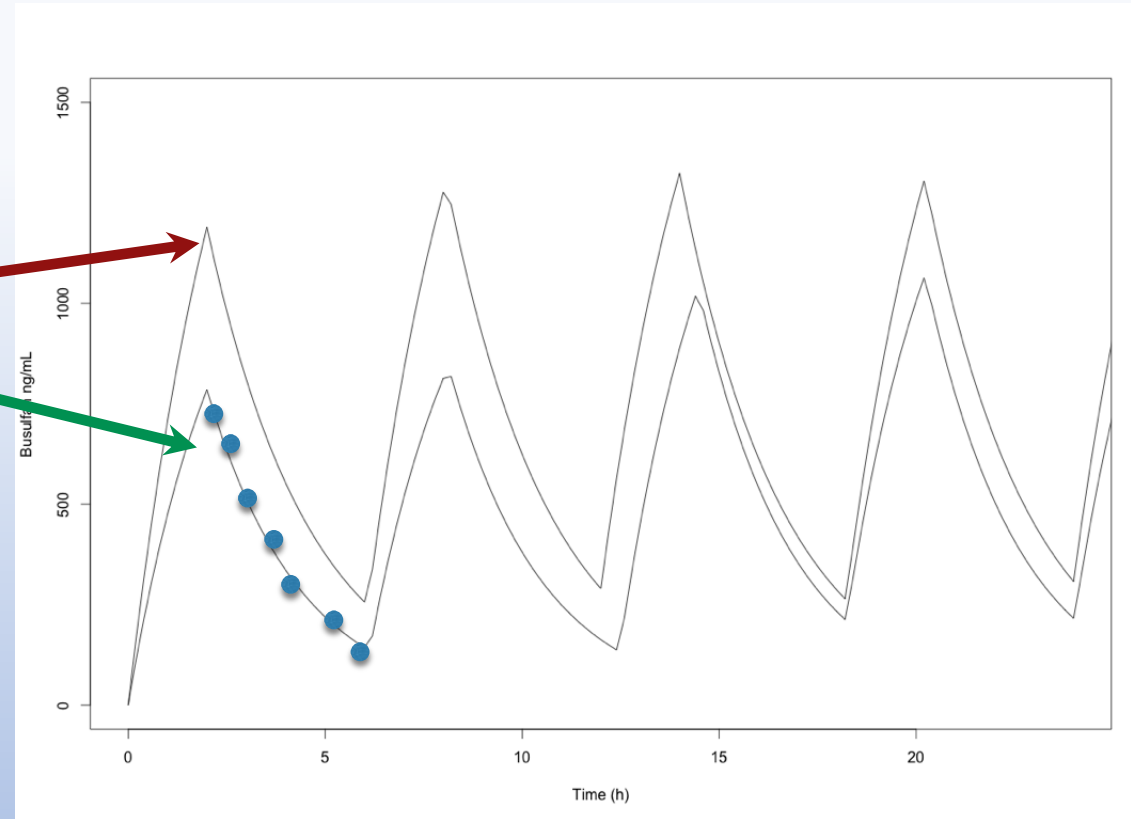
General Workflow

NPAG Start

Iteratively Choose Best Points

Fit the model to the data from subjects

```
PM_fit$run(model, data, …)
```

End

A non-parametric population model
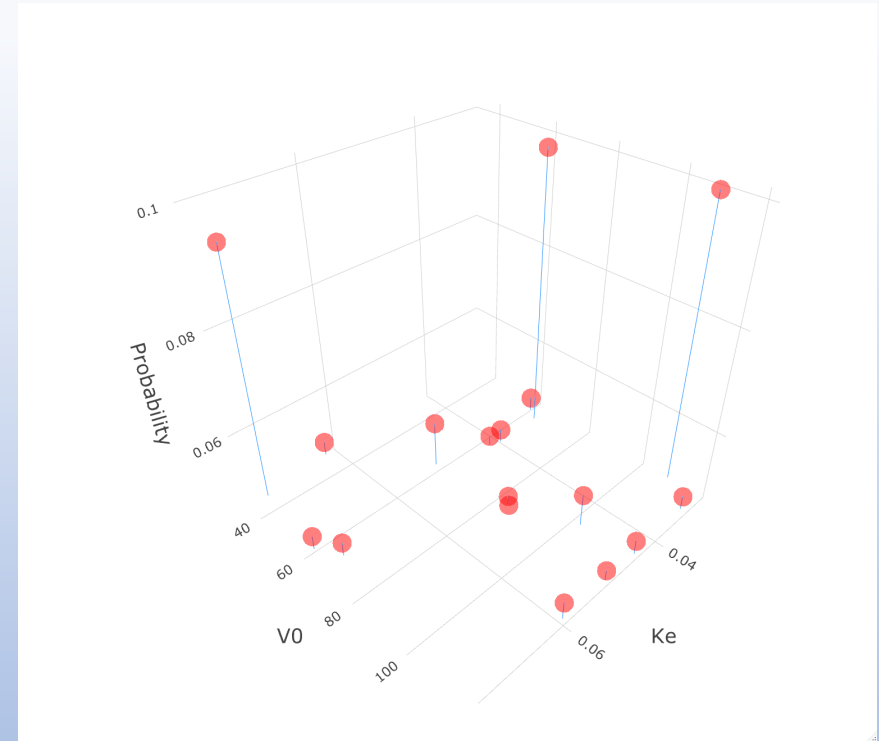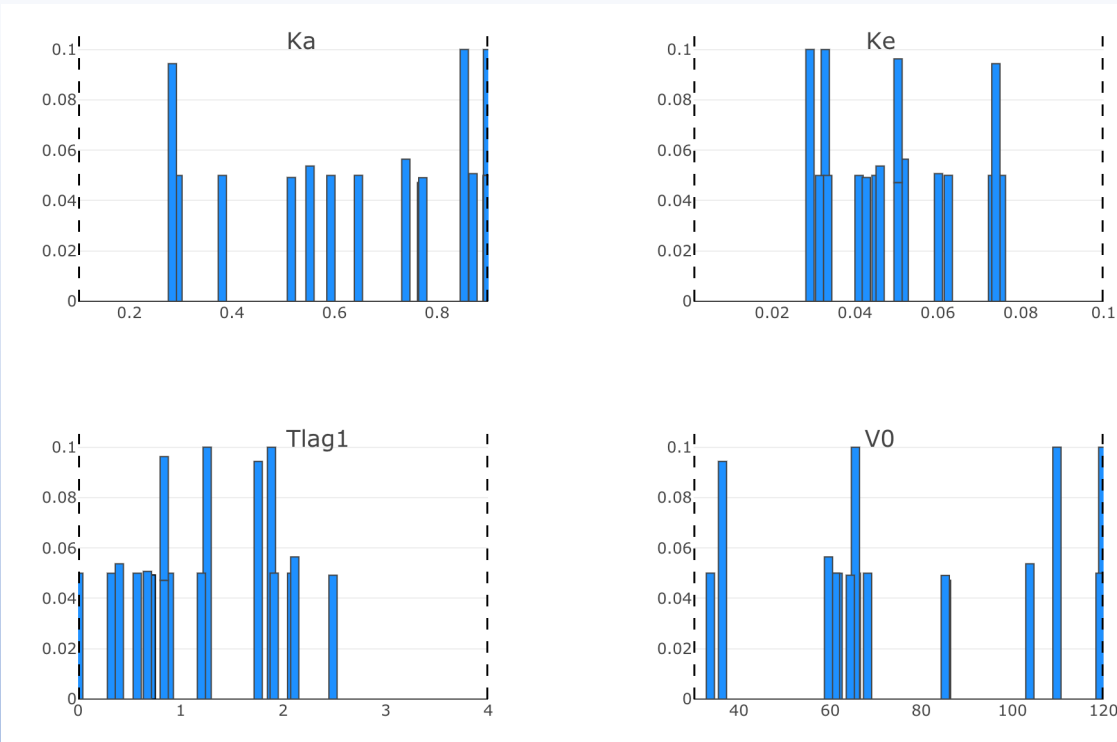
run1$finalplot()

run1$final$plot(V0~Ke)

NPAG runs

```
PM_fit$new(model = "model.txt", data = "data.csv")


PM_fit$run(engine = "NPAG", run, include, exclude,  ode = -4, tol = 0.01, salt, cycles = 100, indpts, icen = "median",  aucint, idelta = 12, prior, overwrite = F, nocheck = F, parallel = NA)
```

Model

- A PM_model object or the quoted name of a model file in your working directory, e.g `model="model2.txt"`

- If your model filename is the default, "model.txt", you can omit

Data

- The name of a PM_data object or the quoted name of a data file in your working directory, e.g `data="data2.csv"`

- If your data filename is the default, "data.csv", you can omit

PM_fit$run(engine)

- Default is "NPAG" but could be "IT2B"

PM_fit$run(run)

- Optional argument not normally needed
- Specify as a number to force a run to be numbered as such, e.g. `run=4` will result in a folder labeled "4" with the results of the run
  - If the folder specified by `run` exists already, for safety, you will see the following error if `overwrite` is not set to TRUE

```
Error:
 4  exists already.  Set overwrite=T to
overwrite.
```

PM_fit$run(include)

- Optional vector of subject IDs to include in the run
- If all IDs are numeric, does not need to be quoted, but should be quoted for alpha IDs
- e.g. `include=c(1:4,8,9,12)` or `include=c("A1001", "B297")`

PM_fit$run(exclude)

- Optional vector of subject IDs to exclude from the run
- If all IDs are numeric, does not need to be quoted, but should be quoted for alpha IDs
- e.g. `exclude=c(1:4,8,9,12)` or `exclude=c("A1001", "B297")`
- You cannot use include and exclude together

PM_fit$run(ode)

- Sets the "stiffness" of the ordinary differential equation solver, with more stiffness taking longer

- Ignored for algebraic models

- Specified as a log value, e.g. -4 (default)

- Not usually required. Higher numbers will result in faster runs but possibly less accurate log-likelihoods and slightly different parameter value distributions

PM_fit$run(tol)

- Controls the convergence criterion for NPAG.
- Smaller numbers make it harder to converge.
- Default is 0.01

PM_fit$run(salt)

- This is different than bioavailability. It is a vector of numbers for the salt values or active fraction of each input, e.g. `salt=c(0.8, 1)`

- Default is 1 for all inputs

- Sometimes is other than 1, e.g. aminophylline salt value is 0.8

PM_fit$run(cycles)

- Control the maximum number of cycles for NPAG to run, e.g. `cycles=5000`.

- Default is 100

- Set to 0 in combination with a specified `prior` to calculate Bayesian posteriors for external data

- The most common argument to change, other than `model` or `data`.

PM_fit$run(indpts)

- The index of the number of starting points
- Automatically set by Pmetrics based on number of random parameters in #PRI block of model file

Npar	Index	Gridpoints
2	1	2129
	2	5003
3	3	10007
4	4	20011
	5	40009
5	6	80021
6+	101, 102, …108	(100-index)*80021

PM_fit$run(icen)

- The summary function used to generate predictions in the HTML report at the end of a run

- Default is "median", but could be "mean"

- Both sets of predictions are available in results loaded by `PM_load()`

PM_fit$run(aucint)

- Old argument maintained for backwards compatibility
- Determined AUC interval to be automatically included
- `makeAUC()` has replaced this argument

PM_fit$run(idelta)

- Controls the frequency for Bayesian prior/posterior predictions for each subject (i.e. PMpost and PMpop objects)

- Specify in 1/60 time units, typically minutes (be careful if your time units are days or something else)

- Default is 12, e.g. 5 predictions per hour

PM_fit$run(prior)

- Specify a previous run to use as a prior for the current run, e.g. `prior=3`

- Could be specified as a filename of a prior density file placed into the /Runs folder at the time of a run
  - Called "DEN0001" and found in the /outputs folder of the run that serves as the prior

PM_fit$run(overwrite)

- Overwrite a previous run
- Used in combination with `run` argument

PM_fit$run(nocheck)

- Suppresses `PMcheck()` which automatically runs silently with each `NPrun()`.

- Default is FALSE.

PM_fit$run(parallel)

- Parallel mode will be selected automatically for models that have a #DIF block, i.e. use differential equations

- Serial mode will be selected for algebraic models

- Use this argument to override the default behavior

NPAG examples

- Run 100 cycles with model.txt and data.csv

  - `data1 <- PM_data$new("data.csv"); fit1 <- PM_fit$new(data); fit1$run()`

- Run 1000 cycles with model2.txt and data from run 1

  - `fit2 <- PM_fit$new(data1, "model2.txt"); fit2$run(cycles=1000)`

- Run 5000 more cycles from run 3, which did not converge

  - `fit2$run(prior = 2, cycles = 5000)`

- Use model in run 5 to calculate Bayesian posterior parameters in new dataset, data2.csv

  - `fit5$run(prior=5,data="data2.csv",cycles=0)`

IT2B runs

- **IT**erative **2**-stage **B**ayesian
- Parametric algorithm
- Arguments are the same as for NPAG except no `indpts`, `aucint`, `idelta`, or `prior`.

ERRrun

- Variation on IT2B

- Designed to estimate C0, C1, C2, C3 directly from data
  - Least preferred method as estimates are model dependent

- Arguments are the same as for IT2B with addition of `search="cursory|medium|extensive"` for depth of coefficient search

Comparing Runs

- PM_compare(x, y, ..., icen = "median", outeq = 1, plot = F)

- **x** = PM_result1, **y** = PM_result2, **...** = additional run numbers or arguments to `plot.PMop()` if `plot=T`

- **icen** = basis for predictions, either "median" (default) or "mean"

- **outeq** = output equation to compare

- **plot** = plot obs vs. pred or not

Comparing runs

	run	type	nsub	nvar	par	converge	-2*LL	aic	bic	popBias	popImp	popPer_RMSE	postBias	postImp	postPer_RMSE	pval
1	1	NPAG	20	4	Ka Ke V Tlag1	FALSE	440.0	450.4	464.6	1.217	18.31	46.17	-0.08785	0.9888	12.18	NA
2	2	NPAG	20	4	Ka Ke V0 Tlag1	FALSE	439.7	450.2	464.4	1.424	17.44	43.07	-0.08622	1.0000	12.39	0.607
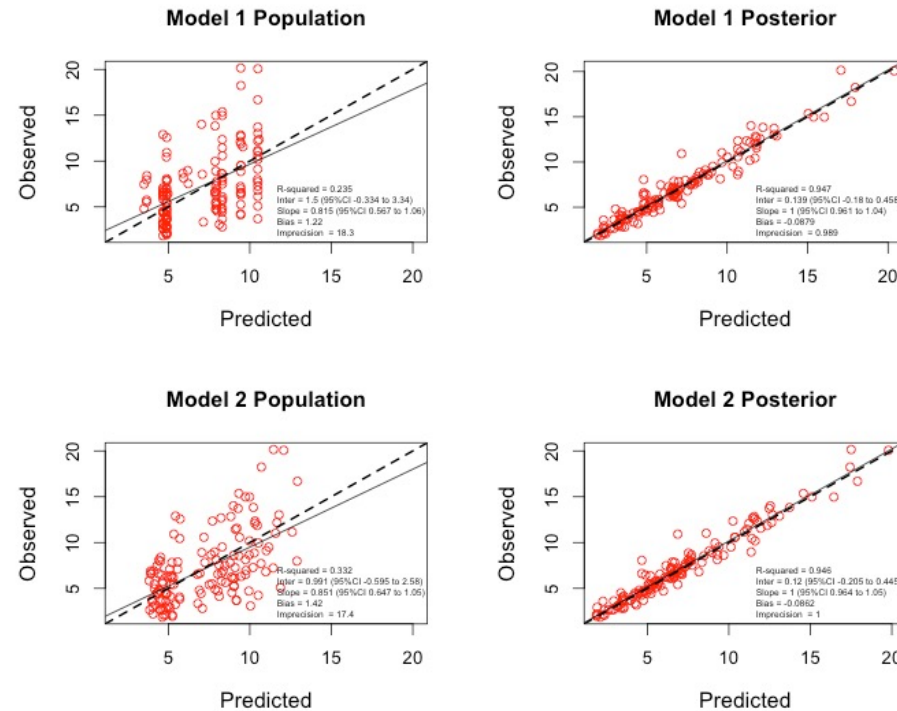

```
PM_compare(run1,run2,plot=T,cex.stat=0.9)
```