Accurate Interpretation of the Visual Predictive Check in order to Evaluate Model Performance

T.M. Post[1], J.I. Freijer[1], W. de Winter[1], M. Danhof[1,2], B.A. Ploeger[1,2]

[1] LAP&P Consultants BV, Leiden, The Netherlands, email: info@lapp.nl

[2] Leiden University, Leiden / Amsterdam Center for Drugs Research, Leiden, The Netherlands.

Introduction

A valuable method to characterize model performance is the Visual Predictive Check (VPC) [1,2]. The purpose is to determine whether a model can reproduce the variability in the observed data. However, it solely relies on subjective graphical inspection of the distribution in the simulated *versus* the observed data [2,3]. It is not evaluated whether the expected random distribution of the observations around the predicted median trend is realized. Moreover, it does not account for the number of observations at each time-point or the influence and information residing in missing data (e.g. below LOQ and dropout in longitudinal studies) [4, 5, 6]. As a result, the model fit might be perceived as being biased, whereas this could also be due to an unbalanced distribution of the observations over time. Therefore, we propose a method for a more accurate and objective interpretation of model performance using the Visual Predictive Check, taking into account the amount of observed data and the influence of missing data.

Methods

With the VPC, the 5th, 50th and 95th percentiles are calculated from the results of 1000 simulations with the optimized model and model parameters. In the proposed extension to the VPC, the following steps are added:

i) the percentage of observations above and below the model predicted median (50th percentile) at each time-point is calculated and visualized, as well as the amount of missing data at each time-point based on the expected number of observations. The median of the observed data is calculated as: (percentage above + below model predicted median) / 2 (**QVPC**).

ii) the 5th, 50th and 95th percentiles of the bootstrapped median of the original observations at each time-point, accounting for the number and assumed position of missing data (informative [above, below] or non-informative), is compared to the model predicted median (**BVPC**).

The method is illustrated by two examples; a simulated PK study (20 subjects) and a phase III PD study (1204 subjects) [7]. First, PK data is generated and fitted with the PK model. A standard VPC is performed with in addition a QVPC and a BVPC, to clarify the current approach. Subsequently, the amount of data is decreased in order to exemplify the current approach and to illustrate the influence of data below LOQ on the interpretation of model performance. The PD example then illustrates how the effect of missing data on the predictive performance can be evaluated with this approach.

BVPC with S-PLUS© code

1. Obtain median statistics of available observations at each time-point and resample in *Bootsamp.* [NoBS: number of bootstrap replications; 1000]

```
Indices <- resamp.get.indices(bootstrap(x,median(x),seed=5,B=NoBS,save.indices=T))
BootSamp <- matrix(nrow=length(x),ncol=NoBS)
BootSamp[,] <- x[Indices[,]]
```

2. Determine extreme observations for available observations at each time-point and the number of missing observations [NoM: number of missing observations]. Fill *Dropoutmatrix* with simulated data, based on the assumption whether the missing data is above or below predicted median [AsM: probability of data being above predicted median]. Combine *Bootsamp* and *Dropoutmatrix.*

```
maxx <- max(x) minx <- min(x)
Dropoutmatrix <- matrix(nrow=NoM,ncol=NoBS)
if (NoM>=1){for(j in 1:NoBS){
      for (i in 1:NoM)
      {     yes <- rbinom(1,1,AsM)
            Dropoutmatrix[i,j] <- maxx*yes + minx*(1-yes)}
      }
            Both <- rbind(BootSamp,Dropoutmatrix)
      }else{
            Both <- BootSamp
      }
```

3. Determine the median for each replicate dataset in *Both* and determine the 5th, 50th, 95th percentiles for the total of the medians, only if missing data comprises less than 50% of the expected amount of data at each time-point.

REFERENCES

[1] Y. Yano, S.L. Beal, L.B. Sheiner, Evaluating pharmacokinetic/pharmacodynamic models using the Posterior Predictive Check, J Pharmacokinet Pharmacodyn 2001; 28(2) 171-192

[2] N. Holford, The Visual Predictive Check – Superiority to Standard Diagnostic (Rorschach) Plots, PAGE 14 (2005) Abstr 738 [www.page-meeting.org/?abstract=738]

[3] P.R. Jadhav and J.V.S. Gobburu, A New Equivalence Based Metric for Predictive Check to Qualify Mixed-Effects Models, AAPS Journal 2005; 7(3): E523-E531

[4] C. Hu, M.E. Sale, A joint model for nonlinear longitudinal data with informative dropout, J Pharmacokinet Pharmacodyn 2003; 30(1) 83-103

[5] L.B. Sheiner, S.L. Beal, A. Dunne, Analysis of nonrandomly censored ordered categorical longitudinal data from analgesic trials, J Amer Stats Assn 1997; 92 1235-1255

[6] J.P. Hing, S.G. Woolfrey, D. Greenslade, P.M.C. Wright, Analysis of toxicokinetic data using NONMEM: impact of quantification limit and replacement strategies for censored data, J Pharmacokinet Pharmacodyn 2001; 28(5) 465-479

[7] W. de Winter, J. DeJongh, T. Post *et al.*, A mechanism-based disease model for comparison of long-term effects of pioglitazone, metformin and gliclazide on disease processes underlying type 2 diabetes mellitus, J Pharmacokinet Pharmacodyn, 2006

Results

PK

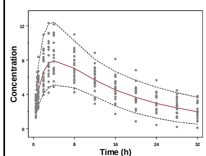

Fig 1a. VPC
Grey dots = observed data
Red solid line = model predicted median
Dashed black lines = 5th and 95th percentile

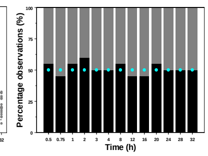

Fig 1b. QVPC
Dark grey bar = observations > predicted median
Black bar = observations < predicted median
Cyan dots = median of available observed data

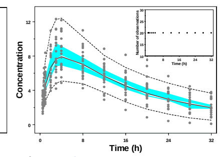

Fig 1c. BVPC
Cyan area = range (5-95th percentiles) bootstrap median
White solid line = 50th percentile bootstrap median

> VPC: **This approach relies on graphical inspection of the simulated ranges *versus* the perceived distribution of the observed data, without accounting for the distribution of the data or the amount of missing data (VPC; Fig 1a, 2a, 3a). Bias may be perceived in the model fits due to the amount of observed data (Fig 2a) or the distribution (skewness) of the data (Fig 3a).**

> QVPC: **Deviation from 50% of the median based solely on the *available* observations presents the uncertainty in this statistic due to missing data (QVPC; Fig 1b, 2b, 3b). Additionally, these plots objectively reflect the position and amount of data regardless of their density (either 20 [Fig 2b] or 1204 [Fig 3b] subjects) around the model predicted median.**

> BVPC: **Model performance can be judged from the position of the model predicted median relative to the range of the bootstrapped median (BVPC; Fig 1c, 2c, 3c). This diagnostic also reflects whether it is possible to judge model performance at a certain time-point as it takes account of the amount of data relative to the expected amount of data (Fig 2c).**

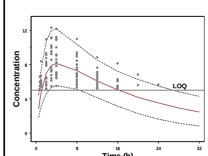

Fig 2a. VPC

Grey dots = observed data
Red solid line = model predicted median
Dashed black lines = 5th and 95th percentile

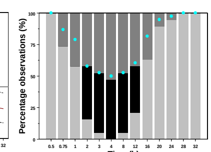

Fig 2b. QVPC
Cyan dots deviating from 50% presents uncertainty in observed data median.

Dark grey bar = observations > predicted median
Black bar = observations < predicted median
Cyan dots = median of available observed data
Light grey bar = missing observations

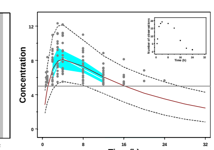

Fig 2c. BVPC
Assumption: missing observations are below model predicted median. For missing > 50% of observation no bootstrap median is calculated.

Cyan area = range (5-95th percentiles) bootstrap median
White solid line = 50th percentile bootstrap median

PD

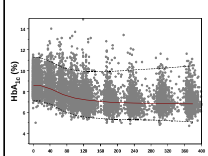

Fig 3a. VPC

Grey dots = observed data
Red solid line = model predicted median
Dashed black lines = 5th and 95th percentile

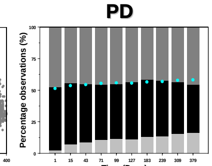

Fig 3b. QVPC
Cyan dots deviating from 50% presents uncertainty in observed data median.

Dark grey bar = observations > predicted median
Black bar = observations < predicted median
Cyan dots = median of available observed data
Light grey bar = missing observations

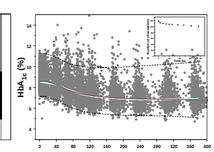

Fig 3c. BVPC
Assumption: missing observations random around model predicted median.

Cyan area = range (5-95th percentiles) bootstrap median
White solid line = 50th percentile bootstrap median

Additional relevant application of the QVPC (and BVPC):
> diagnostics for random and non-random missing

Subjects dropping out of the study were accounted for in analogy to the method described by Hu and Sale [4]; The value of the observation on the occasion previous to dropout was identified. The position (e.g. above or below) of this value was compared to the model predicted median at the corresponding time-point. For the remainder of the study, the subject was then assigned that position. In this manner a cumulative percentage of dropout-subjects over time was visualised in the QVPC. Other missing data were randomly assigned above or below the model predicted median.

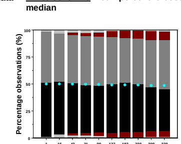

Fig 4. QVPC
Visualising random missing data and possible non-random missing (informative dropout) [4].

Dark grey bar = observations > predicted median
Black bar = observations < predicted median
Cyan dots = median of available observed data
Light grey bar = random missing observations
Dark red bar = missing observations due to dropout

Conclusion

The proposed method facilitated the evaluation of model performance by the linking the VPC to the observed data while accounting for the amount of observed data and the influence of missing data. The applied method puts the VPC in perspective in relation to the distribution of the observations, regardless the density of the data. As a result, this leads to a more accurate and objective evaluation of model performance.

VPC: points to account for

- ✓**Amount of available and missing observations at each time-point**
- ✓**Distribution of the observations around the model predicted median**
- ✓**Uncertainty in the median of the available observations**
- ✓**Compare model predicted median to the range of the bootstrapped median**