The handle http://hdl.handle.net/1887/35437 holds various files of this Leiden University dissertation

**Author:** Bellanti, Francesco  
**Title:** From data to models : reducing uncertainty in benefit-risk assessment : application to chronic iron overload in children  
**Issue Date:** 2015-09-24
CHAPTER 4

Sampling optimisation in pharmacokinetic bridging studies: use of deferiprone in children with β-thalassaemia

Francesco Bellanti, Vincenzo Luca Di Iorio, Meindert Danhof, and Oscar Della Pasqua

Submitted for publication

Summary
Despite a wide experience with deferiprone, the optimum dosage in children aged less than 6 years remains to be established. This analysis is aimed at optimizing the design of a prospective clinical investigation for the evaluation of deferiprone pharmacokinetics in children. A one-compartment model with first order oral absorption was used for the design optimization. Different sampling schemes were evaluated under the assumption of a constrained population size. A sampling scheme with 5 samples post-dose per subject was found to be sufficient to ensure accurate characterization of the pharmacokinetics of deferiprone. Whereas the accuracy of parameters estimates was high, precision was slightly reduced due to the small sample size (> 30% for Vd/F and KA). AUC values (mean and SD) were found to be 33.37 (19.24) and 35.61 (20.22) μg/ml.h and Cmax values 10.17 (6.05) and 10.94 (6.68) μg/ml in sparse and frequent sampling respectively. The results illustrate how ED-optimality concepts can be used to support PK bridging. Predefined sampling schemes and sample sizes do not warrant accurate model structure and parameter identifiability. Of importance is the accurate estimation of the magnitude of the covariate effects, as they may determine the dose recommendation for the population of interest.
4.1 Introduction

Patients with β-thalassaemia and other transfusion-dependent diseases develop iron overload from chronic blood transfusions and require regular iron chelation to prevent potentially fatal iron-related complications. Deferiprone (DFP) is the most extensively studied oral iron chelator to date. It has been authorized in Europe in 1999 for the treatment of iron overload in patients with β-thalassaemia major when deferoxamine (DFO) is contraindicated or inadequate. Despite a wide experience of DFP there are limited experimental data available on DFP in children and no data in children under 6 years of age.

Clinical studies, mostly in patients with beta-thalassaemia, have demonstrated that deferiprone at 75 to 100 mg/kg/day is capable of reducing iron burden in regularly transfused iron-overloaded patients. The degree of iron loading is directly related to the level of iron intake from transfusions. Iron excretion with DFP, like with any chelator, was found to be dose-related. However, factors affecting response to deferiprone appear to include the degree of iron overload and duration, dosage and compliance with therapy. Although few long term comparative data are available, DFP at the recommended dosage of 75 mg/kg/day appears to be non-inferior to deferoxamine in the adult population. However, compliance is superior with DFP.

The optimum dosage of DFP in children less than 6 years of age remains to be established. Given that dose adjustment may be required in children, the aforementioned findings highlight the need for optimizing the dosing regimen and gathering supporting evidence for the dose rationale for subsequent assessment of efficacy in the pediatric population.

The information available so far in the adult population can be used to integrate the lack of knowledge in the pediatric population. The E11 guideline of the International Conference on Harmonization (ICH) supports the use of PK bridging concepts for the development of drugs in the pediatric population. Nevertheless, bridging studies can be implemented only if the following criteria are met: in the populations of interest the medicinal product should have the same indication, the disease process should be similar and the outcome of therapy should comparable. This is true and applicable to patients affected by β-thalassaemia or other transfusion-dependent diseases.

Practical and ethical constraints impose special requirements for clinical trials in children. The application of population pharmacokinetic (PK) analysis and PK bridging to sparse data allows reducing the burden in such a vulnerable population; yet it is important to optimize the quality of the information gathered. The quality of the study can be dramatically improved through design optimization analysis. However, a PK model is needed to apply this methodology. When extrapolating information
from adults to children we have to make use of a hypothetical model which is derived from a different population than the population of interest. ED-optimality concepts can be applied to handle the uncertainty during the optimization procedure. Several studies have already shown how ED-optimization can be successfully applied to the design of clinical studies in children when extrapolating information from the adult population \(^{16-21}\).

Based on simulation scenarios that take into account the impact of developmental growth, the aim of this analysis is to optimize the sampling times for the evaluation of the pharmacokinetics of deferiprone in a prospective clinical investigation in children younger than 6 years of age. The results of this trial will be subsequently used to define the most appropriate dosing regimen for this population.

4.2 Methods

Prospective Clinical Study: Design
A prospective study has been proposed to establish the pharmacokinetics of deferiprone in children. This will be investigated in a prospective multi-centre, randomized, single blind, and single dose study in patients affected by transfusion dependent heamoglobinopathies aged less than 6 years. Sample size of the study will range between a minimum of 18 up to a maximum of 30 evaluable pediatric patients. Patients will be randomized to three dose levels (8.3, 16.7 and 33.3 mg/kg) and will be exposed to a single dose of deferiprone. A maximum of 5 samples will be collected per patient. An optimization algorithm will be applied to evaluate the best sampling times in order to ensure high precision in parameter estimates and PK model identifiability.

Sampling Times Optimization
Several actions have been taken throughout the analysis that can be summarized in 6 major steps as depicted in the following flow-chart (Figure 1). Each step will be briefly discussed in the following paragraphs.
**Hypothetical PK Model**

A one-compartment PK model with first-order absorption, lag-time to central compartment, and first-order elimination was used for the optimization of the sampling scheme. Between-subject variability (BSV) was estimated for apparent clearance (CL/F), apparent volume of distribution (V/F), and absorption rate constant (Ka). Residual variability was characterized by a proportional error model with a weighting factor.

**Competing Models: extrapolation to the pediatric population**

The reference model was previously developed by our group to explore DFP exposure in adults. With the purpose of optimizing the design of a prospective pediatric trial, the original model has been modified with the inclusion of two different covariate models, namely M1 (body weight linearly correlated with CL/F and V/F), and M2 (fixed allometric scaling: exponent of 0.75 on CL/F and 1.00 on V/F), in order to extrapolate deferiprone exposure to children.

Bearing in mind the objective of extrapolation across populations, focus was given to the model validation steps, which yield information about the variance structure and variance-covariance matrix. Visual predictive check (VPC) and NPDE summaries have been used to validate the model and to assess the suitability for simulation purposes. The software NONMEM (non-linear mixed effect modeling; release version 7.2.0) has been used for the procedure.

**M1**: the inclusion of body weight as a covariate on CL/F and V/F according to linear models was found to give the highest improvement in model performance in the previous investigation. The covariates were not included due to increase in model instability. In this case, given the different objective, we have considered including weight on CL/F and V/F into the final covariate model.
M2: Since the purpose of the analysis is to extrapolate information to the pediatric population, the use of allometric scaling (one of the current standard approach to extrapolate across populations) has been considered to evaluate possible discrepancies in optimizing the sampling schedule with two different approaches. Model parameters have been re-estimated with the new covariate relationships and the model has been tested for simulation purposes.

Diagnostic criteria, such as visual predictive checks and NPDE summaries have been used to assess model performance before the optimization of the study design (ED optimality) would be implemented. For both M1 and M2, Visual predictive check (VPC) plots indicate that model is not biased and is suitable for simulation purposes. In addition, NPDE summaries indicate that the discrepancy between predicted and observed values can be assumed to be normally distributed. VPC and NPDE for both models are provided in Figures S2 to S5 (see Appendix).

Optimization steps (and criteria)
The two hypothetical models have been used to identify the optimal sampling schedule in children after single dose deferiprone. The software for population experimental design “PopED” (release version 2.12) has been used to optimize sampling times and to assess precision in parameter estimates by evaluating the coefficient of variation (CV) for each parameter 23–27. Subsequently, the software NONMEM (non-linear mixed effect modeling; release version 7.2.0), has been used after sampling times optimization in order to assess model stability, and accuracy (RE: relative error) and bias (SME: standard mean error) in parameters estimates.
The following 4 scenarios have been evaluated in PopED to account for possible discrepancies between the two covariate models and differences in sample size:
a) M1 with 18 subjects;
b) M1 with 30 subjects;
c) M2 with 18 subjects; and
d) M2 with 30 subjects.
Sampling times have been optimized in the 4 scenarios for a simulated pediatric population which presented the following demographic characteristics: 50% males and females, and mean body weight of 20.5 kg (SD: 5.4). Subjects have been randomized to 3 dose levels as in the study design described above.

Information gathered through the optimization of the sampling times in PopED has then been used to create seven new realistic scenarios (each consisting of 3 sampling schemes) as
a result of a compromise between full optimization and feasibility in a real clinical trial. The seven scenarios have been compared and evaluated in order to define the final sampling schedule for the PK study. Furthermore, an extra scenario, consisting of an empirical, non-optimized sampling scheme has been evaluated along with the previous seven. Given that no significant differences have been observed between the original four scenarios in PopED, scenario “a” (M1 with 18 subjects) has been selected and used for the final optimization step; this allowed also reducing significantly the computational effort of the analysis.

4.3 Results

Sampling times optimization in PopED

The results of the optimization of the sampling times is summarized in Figure 2, where the actual sampling times obtained in the 4 scenarios (“a”, “b”, “c”, “d”) are plotted. Each bar represents a sampling time selected during the optimization procedure, whereas each color indicates the contribution (in percentage) of the different scenarios.

![Figure 2. Sampling times obtained by ED-optimality. Sampling times selected during the optimization steps using ED-optimality: in red, green, dark blue and light blue are shown the time selected for scenarios a, b, c and d respectively. Percent of total indicates the percentage of cases for each set of optimized sampling times generated by PopED.](image)
The data suggest that, independently of model and number of subjects, approximately 92% of sampling times should be collected within three intervals, namely:

- time window A: 32% in the range 10 to 20 minutes;
- time window B: 37% between 40 and 80 minutes;
- time window C: 23.75% after 200 minutes;

Based on the information collected in the previous step, and bearing in mind the compromise between full optimization and feasibility in a real clinical trial, seven sampling schemes (shown in Table 1 as scheme 1 to 7) have been generated and subsequently evaluated in PopED and NONMEM.

**Table 1.** Scenarios evaluated for sampling scheme selection.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Sampling schemes</th>
<th>Scenario</th>
<th>Sampling schemes</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Optimal Design</td>
<td>5</td>
<td>a: 10, 25, 50, 70, 360 b: 15, 45, 60, 270, 420 c: 20, 55, 75, 330, 480</td>
</tr>
<tr>
<td>1</td>
<td>a: 10, 25, 45, 70, 360 b: 15, 40, 60, 180, 420 c: 20, 55, 75, 240, 480</td>
<td>6</td>
<td>a: 10, 40, 65, 85, 360 b: 15, 45, 60, 270, 420 c: 20, 55, 75, 330, 480</td>
</tr>
<tr>
<td>2</td>
<td>a: 10, 30, 45, 180, 360 b: 15, 40, 60, 240, 420 c: 20, 50, 75, 300, 480</td>
<td>7</td>
<td>a: 10, 40, 65, 85, 360 b: 15, 50, 70, 270, 420 c: 20, 55, 75, 330, 480</td>
</tr>
<tr>
<td>3</td>
<td>a: 10, 40, 70, 180, 360 b: 15, 50, 80, 240, 420 c: 20, 60, 90, 300, 480</td>
<td>8*</td>
<td>30, 60, 120, 240, 480</td>
</tr>
<tr>
<td>4</td>
<td>a: 10, 25, 50, 75, 330 b: 15, 45, 60, 240, 360 c: 20, 55, 75, 270, 420</td>
<td>/</td>
<td>/</td>
</tr>
</tbody>
</table>

* Empirical sampling scheme reflecting the current practice, i.e., non-optimised design.

**Evaluation of seven realistic sampling schemes**

As previously mentioned in the methods section, given that no major differences have been observed in the previous step between the two different models (M1 and M2) and different number of subjects (18 vs. 30), only model M1 with a total number of 18 subjects was used for the second part of the optimization.

Table S1 (see Appendix) shows the coefficient of variation (CV) for the different scenarios compared to the optimal sampling scheme. No major differences can be observed between the different scenarios, except for the non-optimized scheme (number 8) in which a remarkably higher uncertainty for the following parameters can be observed: CL slope, V
slope, and Ka. Furthermore, results clearly highlight the poor performance of the model if an empirical (non-optimized) pharmacokinetic sampling scheme is used.

Table S2 (see Appendix) shows model stability results based on NONMEM stochastic simulation and estimation (SSE). In this overview, schemes 3 and 7 show higher stability, as compared to the other sampling schemes. On the other hand, scheme 8 shows the worst result out of the 9 scenarios. Finally, Figures 3, 4 and S1 (for Figure S1 see Appendix) show measures of accuracy (RE) and bias (SME) for the main parameters of interest for the different sampling schemes.

Altogether, scheme 7 was the closest one to the fully optimized sampling scheme, providing the best combination of results in terms of bias (SME) and accuracy (RE) of parameters estimates. This was also true in terms of model robustness, with only 1 failed minimization and 435 successful covariate steps out of 500 runs (Table S2; see Appendix).
Figure 3. SME and RE for the slope parameters describing the effect of body weight on clearance and volume of distribution. Standard Mean Error (SME) and Relative Error (RE) estimates indicate, respectively, the bias and accuracy in parameter estimates for the different sampling schemes. Top panels: summary for the slope parameter describing the effect of body weight on clearance. Bottom panels: summary for the slope parameter describing the effect of body weight on volume of distribution.

Figure 4. SME and RE for the volume of distribution in males and females. Standard Mean Error (SME) and Relative Error (RE) estimates indicate, respectively, the bias and accuracy in parameter estimates for the different sampling schemes. Top panels: summary for volume of distribution in males. Bottom panels: summary for volume of distribution in females.
Secondary PK parameters: sparse sampling vs. rich sampling
To further assess the suitability of scheme seven for the prospective PK study, focus was
given to the ability of the model to predict secondary PK parameters. Model-predicted AUC and Cmax based on the sparse sampling scheme were compared with estimates obtained according to a rich sampling scheme (12 samples per subject). AUC has been calculated with
the trapezoidal rule.

Figure 5. AUC and CMAX estimation using sparse (5) vs. frequent (12) sampling. Model predicted
deferoxamine systemic exposure expressed as AUC and Cmax. The final scheme selected during the
optimization procedure (scheme 7) with sparse sampling (5 samples) is compared with rich sampling
(12 samples). Top panels: histogram and boxplot describing the distribution of Cmax. Bottom panels:
histogram of the distribution of AUC. Percent of total represents the percentage of cases for each set
of 500 simulations with 18 patients in each simulated trial. Red: sparse sampling; Blue: rich sampling;
Green: overlapping area.
As shown by the distributions and box-plots of AUC and Cmax (Figure 5), no significant differences were detected for sparse and frequent sampling. AUC values (mean and standard deviation) were found to be 33.37 (19.24) and 35.61 (20.22) μg/ml.h and Cmax values 10.17 (6.05) and 10.94 (6.68) μg/ml in sparse and frequent sampling respectively.

4.4 Discussion and Conclusion

Clinical trials in children represent a very challenging phase in drug development. Given the ethical and practical constraints imposed to experimental protocols in this vulnerable population \(^{12-15}\), the information gathered per subject becomes significantly more important, as compared to the adult population. In addition, increasing evidence suggests the lack of suitability of empirical protocols in pediatric research. This limitation has therefore prompted clinical scientists and drug developers to consider the use of alternative approaches for the evaluation of pharmacokinetics, efficacy and safety in children \(^{28-34}\).

Our results show that highly informative data can be generated whilst reducing the burden of clinical trials in children. From a methodological perspective, these data also reinforce the benefits associated with the use of ED-optimality concepts for the design of pediatric clinical studies \(^{16-21}\). In fact, it should be noted that whereas optimal design is normally based on a model representative of the population of interest, our analysis was aimed at an extrapolation model derived from data available in adult patients, i.e., we have used a hypothetical model. Consequently, the optimization procedures carry a certain degree of uncertainty. Nevertheless, the major advantage of using ED-optimization is that this methodology accounts for the uncertainty parameter estimates and in the effect of covariates during the optimization procedure. More specifically, two scenarios with 20% and 40% uncertainty on the main parameters of interest (clearance and volume of distribution) were evaluated. The data shown throughout this report reflect the first case only. Increased uncertainty in parameter distributions had no significant impact on the optimization and selection of sampling times. Lastly, to ensure a comprehensive evaluation of the assumptions underlying the nature and magnitude of the covariate effects on the systemic exposure in children, we have resorted to two competing models (M1 and M2). These models enabled the identification of the best sampling scheme for the prospective pharmacokinetic study taking into account potential differences in the disposition of deferiprone in children younger than 6 years of age.

Based on our experience in pediatric clinical pharmacology, fully optimized designs (i.e., individually optimized) are not realistically applicable to pediatric trials. Our objective in this regard was to identify a final sampling scheme that resulted from a compromise between
full optimization and feasibility in a real setting. The three time windows selected during the optimization procedures were found to be independent of model specification and sample size. Moreover, they reflect the requirements for estimating specific pharmacokinetic parameters, i.e., lag-time and Ka (time window A), V/F (time window B), and CL/F (time window C) respectively.

Finally, it should be noted that our analysis clearly highlights the poor performance of an empirical (i.e., non-optimized) sampling scheme (scheme 8), especially when dealing with sparse sampling. By contrast, scenario 7 showed the best option in terms of coefficient of variation, relative errors and standard mean errors as well as for what concerns model stability. In addition, it allowed for correct predictions of AUC and Cmax.

**Potential limitations**

The final decision on the sampling scheme to be used in the prospective study had to take another unknown factor into account, namely the uncertainty around the time at which the peak concentration occurs (Tmax). Keeping in mind that this exercise is based on a model derived from pharmacokinetic data in adult patients, there might be some differences in children below 6 years of age. Such differences may occur despite the quick absorption after administration of the drug as an oral solution. They may also be caused by difficulties in the administration of the drug to the very young children. We have therefore included these considerations in scenarios 3, 6 and 7, but there are no data available at the moment to confirm the assumptions regarding the possible shift of Tmax.

In conclusion, our analysis illustrates and confirms that despite feasibility issues, ED-optimality concepts can be used to optimize study design, particularly with regard to the pediatric population. Predefined sampling schemes and sample sizes do not warrant accurate model structure and parameter identifiability. In addition, it shows that the optimization of study design does not require necessarily the use of the final model for the population of interest; the combination between ED-optimization and the information carried by a hypothetical model is sufficient to significantly increase the quality of the information collected in a prospective clinical trial. Nevertheless, remains of particular importance the accurate estimation of the magnitude of the covariate effects, as they may determine the final dose recommendation for the population of interest.
References


Appendix

Table S1. Coefficient of variation for pharmacokinetic parameters of the model assuming a linear relationship between body weight and CL and Vd. Estimates calculated according to study design including 18 patients and 20% uncertainty in the parameter estimates for CL and V.

<table>
<thead>
<tr>
<th>PARAMETER</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>CL intercept*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CL slope</td>
<td>7.4</td>
<td>7.2</td>
<td>6.9</td>
<td>6.9</td>
<td>7.0</td>
<td>7.0</td>
<td>7.0</td>
<td>7.0</td>
<td>29.4</td>
</tr>
<tr>
<td>BSV CL</td>
<td>59.7</td>
<td>57.3</td>
<td>53.9</td>
<td>53.2</td>
<td>55.2</td>
<td>55.4</td>
<td>55.0</td>
<td>54.8</td>
<td>52.7</td>
</tr>
<tr>
<td>V Males</td>
<td>22.5</td>
<td>23.4</td>
<td>24.1</td>
<td>24.1</td>
<td>23.6</td>
<td>23.4</td>
<td>23.2</td>
<td>23.3</td>
<td>23.8</td>
</tr>
<tr>
<td>V Females</td>
<td>43.8</td>
<td>44.6</td>
<td>45.8</td>
<td>45.8</td>
<td>44.9</td>
<td>44.5</td>
<td>44.2</td>
<td>44.3</td>
<td>45.5</td>
</tr>
<tr>
<td>V slope</td>
<td>48.9</td>
<td>49.7</td>
<td>50.8</td>
<td>51.1</td>
<td>49.9</td>
<td>49.6</td>
<td>49.4</td>
<td>49.6</td>
<td>77.4</td>
</tr>
<tr>
<td>BSV V</td>
<td>57.8</td>
<td>59.5</td>
<td>61.6</td>
<td>62.8</td>
<td>60.3</td>
<td>59.5</td>
<td>59.0</td>
<td>59.4</td>
<td>59.9</td>
</tr>
<tr>
<td>Ka</td>
<td>26.9</td>
<td>29.7</td>
<td>33.3</td>
<td>37.3</td>
<td>29.7</td>
<td>29.5</td>
<td>37.0</td>
<td>38.1</td>
<td>379.2</td>
</tr>
<tr>
<td>BSV Ka</td>
<td>37.4</td>
<td>39.2</td>
<td>40.1</td>
<td>40.4</td>
<td>39.2</td>
<td>39.2</td>
<td>40.4</td>
<td>40.4</td>
<td>72.2</td>
</tr>
</tbody>
</table>

*The parameter describing the intercept for the linear function between body weight and clearance has been fixed during the analysis. Therefore no CV values are reported.

Table S2. Evaluation of model stability and parameter identifiability based on different sampling schemes. Values represent the results of 500 runs.

<table>
<thead>
<tr>
<th>SAMPLING SCHEME</th>
<th>Minimisation successful</th>
<th>Covariate Step Successful</th>
<th>Estimate near boundary</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>496</td>
<td>391</td>
<td>108</td>
</tr>
<tr>
<td>1</td>
<td>409</td>
<td>306</td>
<td>118</td>
</tr>
<tr>
<td>2</td>
<td>476</td>
<td>398</td>
<td>82</td>
</tr>
<tr>
<td>3</td>
<td>490</td>
<td>443</td>
<td>47</td>
</tr>
<tr>
<td>4</td>
<td>454</td>
<td>362</td>
<td>101</td>
</tr>
<tr>
<td>5</td>
<td>446</td>
<td>419</td>
<td>28</td>
</tr>
<tr>
<td>6</td>
<td>471</td>
<td>419</td>
<td>33</td>
</tr>
<tr>
<td>7</td>
<td>499</td>
<td>435</td>
<td>63</td>
</tr>
<tr>
<td>8</td>
<td>283</td>
<td>217</td>
<td>100</td>
</tr>
</tbody>
</table>
Figure S1. SME and RE for the slope parameters describing the effect of body weight on clearance and volume of distribution. Standard Mean Error (SME) and Relative Error (RE) estimates indicate, respectively, the bias and accuracy in parameter estimates for the different sampling schemes. Top panels: summary for absorption rate constant (Ka). Bottom panels: summary for lag time.
Figure S2. Visual Predictive Check for model M1. Simulated concentration vs. time course profile of deferiprone according to a pharmacokinetic model in which weight is correlated with clearance and volume according to a linear function. Observed data are plotted using open circles; the solid line represents the median of the simulated data; the dashed lines represent the 5th and 95th percentiles of the simulated data.
Normalised prediction distribution errors (NPDE) for model M1. Given that the accuracy of model predictions also depends on the variance structure, special attention was paid to the evaluation of model misspecifications for the random effects. The normalised prediction distribution errors (NPDE) method was applied for an in-depth diagnosis of potentially poor behaviour. Top left: QQ-plot of the distribution of the NPDE versus the theoretical N (0,1) distribution; Top right: Histogram of the distribution of the NPDE, with the density of the standard Gaussian distribution overlaid; Bottom left: NPDE versus time and Bottom right: NPDE versus PRED.
Figure S4. Visual Predictive Check for model M2. Simulated concentration vs. time course profile of deferiprone according to a pharmacokinetic model in which weight is correlated with clearance and volume according to an allometric function. Observed data are plotted using open circles; the solid line represents the median of the simulated data; the dashed lines represent the 5th and 95th percentiles of the simulated data.
Normalised prediction distribution errors (NPDE) for model M2. Given that the accuracy of model predictions also depends on the variance structure, special attention was paid to the evaluation of model misspecifications for the random effects. The normalised prediction distribution errors (NPDE) method was applied for an in-depth diagnosis of potentially poor behaviour. Top left: QQ-plot of the distribution of the NPDE versus the theoretical N (0,1) distribution; Top right: Histogram of the distribution of the NPDE, with the density of the standard Gaussian distribution overlaid; Bottom left: NPDE versus time and Bottom right: NPDE versus PRED.