



Background and Objectives

When evaluating study designs, methods based on Monte Carlo simulations such as Stochastic Simulation and Estimation (SSE) are considered the gold standard. In the SSEs the "true" model and the sampling schedule is used to simulate a number of datasets and the model parameters are re-estimated. This generates a number of parameter vectors which are used to calculate paramater precision metrics such as RMSE, Bias and RSE. Using these parameter metrics for overall comparison of designs can be cumbersome when there are many parameters of interest. Furthermore, these metrics do not include the correlation of the parameters contained within, for example, an empirical variance-covariance matrix (empCOV) from the parameter vectors estimated in the SSEs. The empCOV has other advantages: (1) it can be compared to the theoretically expected covariance matrix predicted by optimal design theory and (2) the comparison can be done using single values like a normalized inverse determinant of the empCOV (equivalent to the the D-criterion). There is however uncertainty in the calculation of the variance-covariance matrix from SSEs (*figure 1*) and the true empCOV can only be achieved when the number of simulations goes towards infinity. In this work this uncertainty in the empCOV and normalized inverse determinant calculation is addressed by generating confidence intervals of the normalized determinant of the inverse empCOV from bootstraps of the SSE parameter vectors.

Methods and Materials

Sampling schedules with 8 samples per individual were optimized in PopED [1-2] for two example models. The first example was based on the warfarin model previously used for software comparison by Nyberg *et al* [3]. The warfarin model was parameterized as shown in Table 1 and the design consisted of 32 individuals in one group and sampling within 0 to 120 hours and a fixed dose of 70 mg. The second example was a sigmoidal EMAX model with concentration increasing linearly with time and parameterized according to table 2. The design contained 100 individuals in one group and samples between 0 and 50 hours. The optimizations were performed using the determinants (D-optimality) of the full and block diagonal FO and FOCE approximated FIM. The residual errors (RUV) had a combined additive and proportional structure with a fixed additive component.

Parameter	Value
CL	0.15
V	8
Ka	1
ω^2_{CL}	0.07
ω^2_V	0.02
ω^2_{Ka}	0.6
$\sigma^2_{add.}$	0.01 FIX
$\sigma^2_{prop.}$	0.01

Table I. Parameterization of the Warfarin model used for optimization.

Parameter	Value
EMAX	100
EC50	20
γ	1
ω^2_{EMAX}	0.0625
ω^2_{EC50}	0.0625
ω^2_{γ}	0.0625
$\sigma^2_{add.}$	25 FIX
$\sigma^2_{prop.}$	0.0025

Table II. Parameterization of the Emax model used for optimization.

The parameters for all designs were re-estimated (FOCEI) with NONMEM 7.2 [4] using 3000 MC simulated datasets in PsN [5] (SSE). For each bootstrap iteration, 3000 parameter vectors were sampled with replacement from the resulting parameter vectors of the SSE and the inverse determinant of the empCOV was calculated and normalized to the number of parameters. Following the bootstrap, the median and the 2.5th and 97.5th percentiles of all 10000 normalized inverse determinants were calculated.

Results

When evaluating the designs for the warfarin example (*figure 2*), the non-bootstrapped normalized inverse determinants from the parameter vectors showed a difference in performance between all the designs with the block-reduced FO design being the best design. However, the 95% confidence intervals of the normalized inverse determinants revealed none of the optimal designs were, in this example, superior to the initial design.

For the EMAX example (*figure 3*), the single non-bootstrapped normalized inverse determinant calculation showed that, although the designs were similar in performance, the FO Full Optimal design is the best design. The evaluation of the designs with confidence intervals of normalized inverse determinants revealed that there was no significant difference in performance between the optimal designs. However, all optimal designs were superior to the initial design.

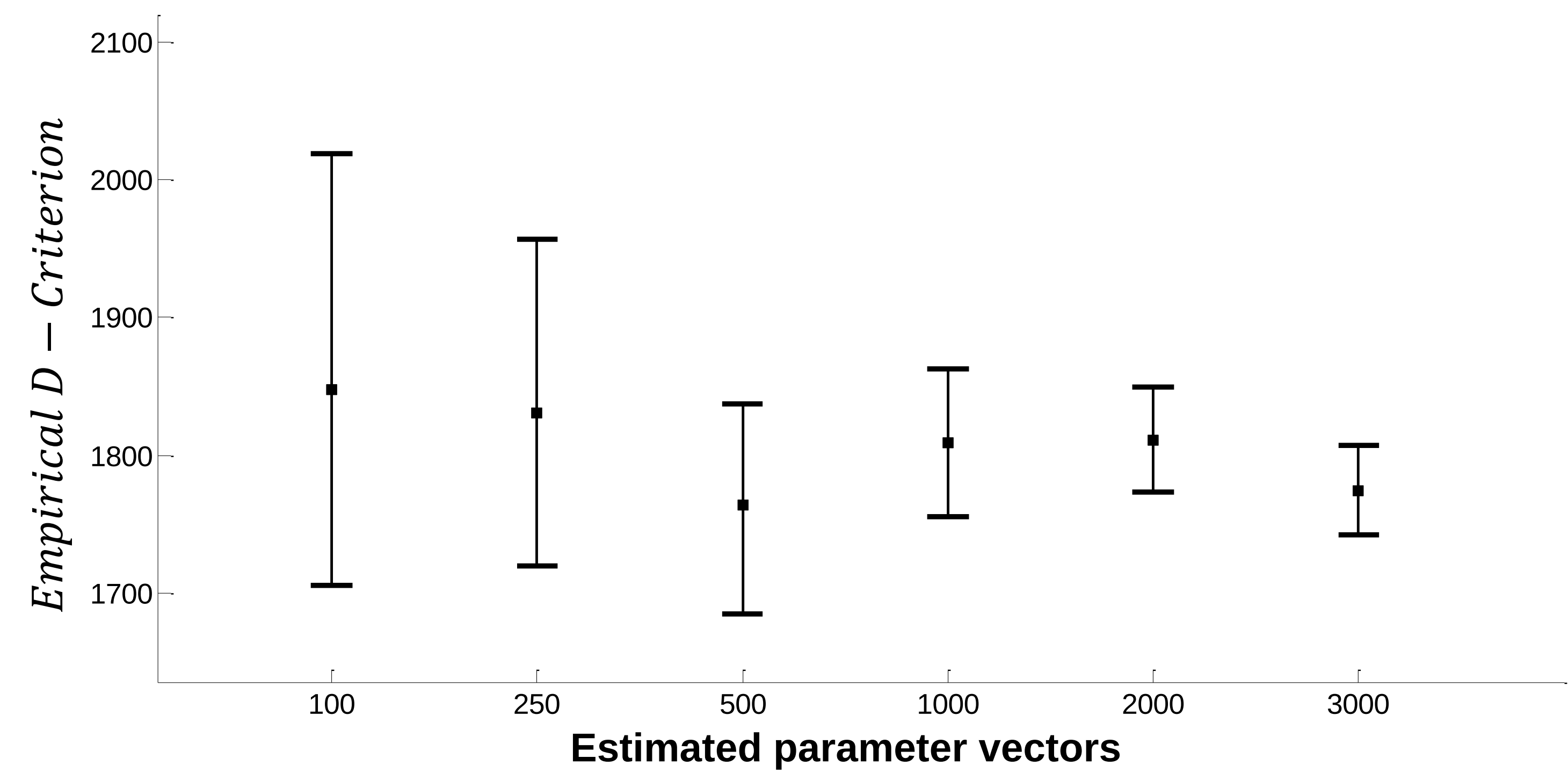


Figure 1: 95% confidence intervals of calculated normalized inverse determinants (empirical D-criterion) of the empirical variance-covariance matrices for the FO approximated Full FIM optimal design in the Warfarin example as the number of simulated datasets and estimated parameter vectors increases. The lower and upper lines represent the 5th and 95th percentiles respectively, the dot represents the median.

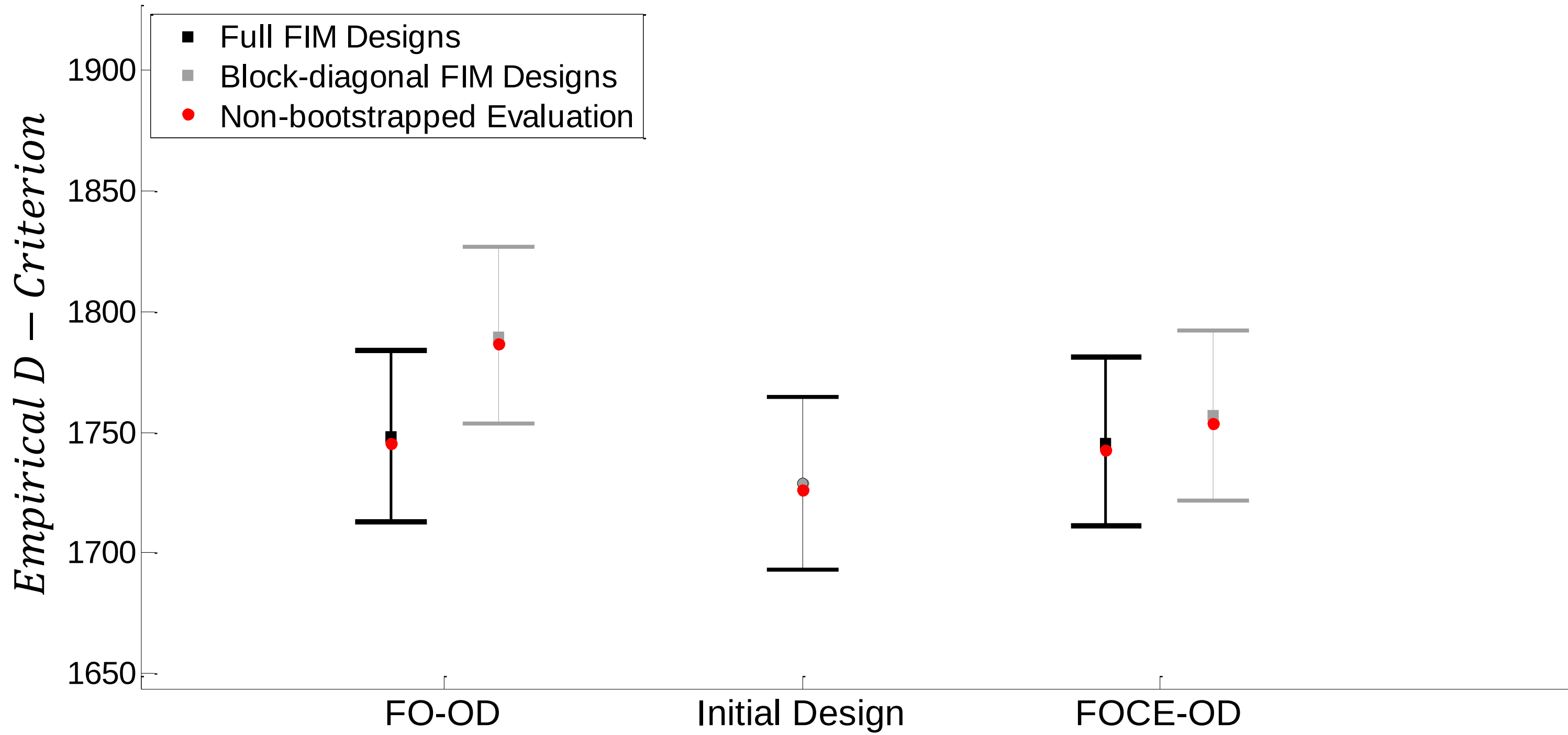


Figure 2: 95% confidence intervals of normalized inverse determinants (Empirical D-Criterion) of the designs in the Warfarin Example. The red dots shows the resulting normalized inverse determinants when calculated from the parameter vectors without performing a bootstrap. The initial design contains the starting sampling schedule for the optimization.

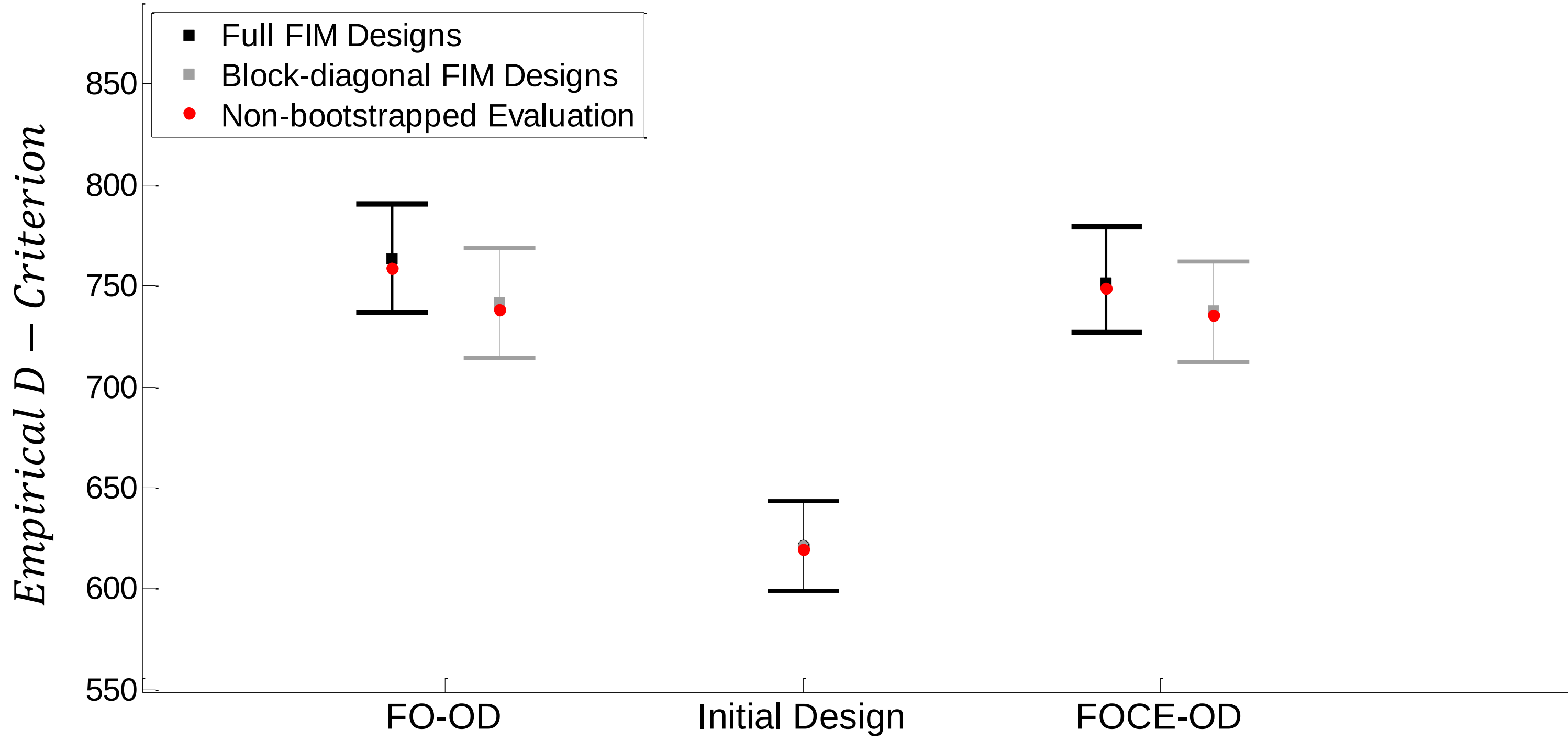


Figure 3: 95% confidence intervals of normalized inverse determinants (Empirical D-criterion) for the designs in the EMAX Example.

Conclusions

- ✓ For these examples, a single calculation of the empirical D-criterion from the SSE vectors would indicate differences in performance between all the optimal designs and the initial design.
- ✓ The bootstrap confidence intervals of the empirical D-criterions show that there is no significant performance difference between most of the designs.
- ✓ Using a single empirical D-criterion from SSE parameter vectors to compare designs may include a risk of false conclusions of design superiority caused by the uncertainty of the calculation. By comparing confidence intervals of empirical D-criterions from a bootstrap procedure, the risk of false conclusions is reduced.

References

[1] Foracchia M, Hooker A, Vicini P, Ruggeri A, "POPED, a software for optimal experiment design in population kinetics.", 2004 Computer Methods and Programs in Biomedicine, 74(1), pp. 29-46

[2] Nyberg J, Ueckert S, Strömberg E.A., Hennig S, Karlsson M.O., Hooker A.C., "PopED: an extended, parallelized, nonlinear mixed effects models optimal design tool.", 2012 Computer Methods and Programs in Biomedicine 108(2), pp. 789-805

[3] Nyberg J, Bazzoli C, Ogungbenro K, Aliev A, Leonov S, Dufull S, Hooker A.C, Mentré F. "Methods and software tool for design evaluation for population pharmacokinetics-pharmacodynamics studies." Br J Clin Pharmacol (2014) 1365-2125

[4] Beal S., Sheiner L.B., Boeckmann, A., & Bauer, R.J., "NONMEM User's Guides." (1989-2009), Icon Development Solutions, Ellicott City, MD, USA, 2009.

[5] Keizer R.J., Karlsson M.O., Hooker A. "Modeling and Simulation Workbench for NONMEM: Tutorial on Pirana, PsN, and Xpose". CPT Pharmacometrics Syst Pharmacol. 2013 Jun 26;2:e50.[<http://psn.sourceforge.net/>]