ARTICLE

Performance of the SAEM and FOCEI Algorithms in the Open-Source, Nonlinear Mixed Effect Modeling Tool nlmixr

Rik Schoemaker[1],* (iD), Matthew Fidler[2] (iD), Christian Laveille[3], Justin J. Wilkins[1] (iD), Richard Hooijmaijers[4], Teun M. Post[4], Mirjam N. Trame[5], Yuan Xiong[6] and Wenping Wang[7]

The free and open-source package nlmixr implements pharmacometric nonlinear mixed effects model parameter estimation in R. It provides a uniform language to define pharmacometric models using ordinary differential equations. Performances of the stochastic approximation expectation-maximization (SAEM) and first order-conditional estimation with interaction (FOCEI) algorithms in nlmixr were compared with those found in the industry standards, Monolix and NONMEM, using the following two scenarios: a simple model fit to 500 sparsely sampled data sets and a range of more complex compartmental models with linear and nonlinear clearance fit to data sets with rich sampling. Estimation results obtained from nlmixr for FOCEI and SAEM matched the corresponding output from NONMEM/FOCEI and Monolix/SAEM closely both in terms of parameter estimates and associated standard errors. These results indicate that nlmixr may provide a viable alternative to existing tools for pharmacometric parameter estimation.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THIS TOPIC?
☑ nlmixr is a free, open-source, nonlinear mixed effect modeling package implemented in R with unknown performance properties.

WHAT QUESTION DID THIS STUDY ADDRESS?
☑ Are the algorithms implemented in nlmixr capable of providing results comparable to NONMEM and Monolix, the industry gold standards?

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?
☑ A near-perfect match between NONMEM/first order-conditional estimation with interaction (FOCEI) and nlmixr/FOCEI, and very high correspondence between Monolix/ stochastic approximation expectation-maximization (SAEM) and nlmixr/SAEM, with occasional suggestions of superiority of the nlmixr implementations, for the scenarios explored.

HOW MIGHT THIS CHANGE DRUG DISCOVERY, DEVELOPMENT, AND/OR THERAPEUTICS?
☑ The availability of a free and open-source modeling tool will allow further rapid uptake of nonlinear mixed effect modeling approaches that have the potential to substantially increase the understanding of drug effects. The availability of a free and open-source tool with robust implementations of cutting-edge algorithms will be of considerable benefit to researchers and students in resource-poor settings.

nlmixr is a freely available, open-source package for R[1] that implements a number of parameter estimation algorithms in the field of nonlinear mixed effect modeling; a stable version is freely available on the Comprehensive R Archive Network (CRAN),[2] and the development version is available from GitHub.[3] The package is primarily intended for the pharmacometric community and allows pharmacometric models to be implemented through the application of ordinary differential equations (ODEs). The pharmacometric accent manifests itself through the availability of fully flexible dosing definitions in terms of the type (e.g., bolus doses or infusions), timing, and number of doses and their amounts, which can vary between and within individuals.

nlmixr builds on RxODE,[4,5] a fast and efficient R package for simulating nonlinear mixed effect models using ODEs with rapid execution as a result of compilation in C. Comprehensive online documentation[6] and an nlmixr tutorial[7] are available.

nlmixr implements a number of parameter estimation algorithms that can be accessed through a common model definition language. These algorithms currently comprise nlme,[8] implemented as a well-established package in R; stochastic approximation expectation-maximization (SAEM);[9] and

[1]Occams Coöperatie UA, Amstelveen, The Netherlands; [2]Novartis Pharmaceuticals Corporation, Fort Worth, Texas, USA; [3]Calvagone, Liergues, France; [4]LAP&P Consultants BV, Leiden, The Netherlands; [5]Novartis Institutes for BioMedical Research, Cambridge, Massachusetts, USA; [6]Certara, Princeton, New Jersey, USA; [7]Novartis Pharmaceuticals Corporation, East Hanover, New Jersey, USA. *Correspondence: Rik Schoemaker (rik.schoemaker@occams.com)

first-order conditional estimation with interaction (FOCEI).[10] Further advances such as implementation of the Gaussian Quadrature algorithm in the common model definition language are in active development. Additional R packages have been developed as add-ons to nlmixr to assess general goodness of fit (xpose.nlmixr[11]) and to provide a project management interface (shinyMixR[12]), with other packages actively developing support for nlmixr (e.g., PharmTeX[13] and ggPMX[14]). Analytical solutions to standard pharmacokinetic models are already implemented for some of the algorithms (nlme, SAEM) and are under active development for others (FOCEI).

For any new tool to be accepted by the pharmacometric modeling and simulation community, it is imperative that the performance characteristics of its estimation algorithms and support functions can be demonstrated to be adequate and comparable to widely used standards.

The primary aim of the current article is to address the question of whether a switch from a standard estimation strategy to an algorithm implemented in nlmixr will produce comparable results. Accordingly, the performances of nlmixr's SAEM and FOCEI estimation algorithms have been compared with implementations in Monolix[15] and NONMEM,[9] respectively, because the implementations in these tools are considered to be the industry standards. Although the nlme algorithm is implemented in nlmixr, it is currently not widely applied in pharmacometrics and has no industry standard implementation as a reference. Consequently, nlme was not investigated in this manuscript.

METHODS

The investigation examines both variations within a model and variations between models. For variations within a model, 500 sparsely sampled data sets are analyzed using the same model. For variations between models, a range of models is applied, each to a distinct richly sampled data set. All data were generated using simulations of trials with four parallel dose levels and 30 subjects per dose level with a total of 120 subjects per trial.

Sparsely sampled data sets fit using a single model
Estimation with sparsely sampled data was investigated for a first-order absorption model with one-compartment disposition and linear elimination. Population values for clearance (CL) of 4.0 L/hour, central volume (Vc) of 70 L, and first-order absorption rate constant (ka) of 1.0 hour^{-1} were used, with 30% interindividual variability (IIV) for all three parameters (implemented as a diagonal matrix with no covariances), and 20% proportional residual variability. Single-dose data for 10,000 subjects were simulated. The population was split into four equal-sized groups that received doses of 10, 30, 60, or 120 mg, and four time points were randomly sampled within the 24 hours after the dose. A total of 500 data sets containing 120 subjects each were resampled from these 10,000 subjects and stratified by dose so that 30 subjects in each resampled data set received one of the four doses using the bootstrap tool of Perl-speaks-NONMEM (PsN).[16] Each resampled data set was then analyzed using the same structural model that was used for simulation using Monolix's SAEM algorithm,

NONMEM's FOCEI algorithm, and nlmixr's SAEM and FOCEI algorithms to allow a paired comparison for each simulated data set of the analysis outcomes.

Richly sampled data sets fit using different models
Richly sampled profiles were simulated for four different dose levels of 30 subjects each, for a range of test models with the following:

- One-compartmental or two-compartmental disposition
- Oral (first-order absorption), intravenous bolus, or intravenous infusion administration
- Linear or Michaelis-Menten clearance.

In addition, the following three dosing and sampling scenarios were investigated:

- A single administration with 19 samples over 72 hours
- Seven repeated daily administrations, with 15 samples over 24 hours after the fourth dose, 19 samples over 72 hours after the seventh dose, and five trough samples
- The single administration profile followed by the repeated administrations profile with a total of 58 samples over 12 days

Full details and all code are provided in the **Supplemental Materials**.

These combinations provided a total of 36 test cases. The IIV was applied to all pharmacokinetic parameters, and all IIVs were set to 30% (implemented as a diagonal matrix with no covariances). Proportional residual variability was set to 20%. All one-compartment models had a population Vc of 70 L, and all two-compartment models had an additional peripheral volume (Vp) of 40 L. For all oral absorption models, ka was set to 1.0 hour^{-1}. All models with linear elimination had a CL of 4.0 L/hour, and for all models with nonlinear Michaelis-Menten elimination, CL was replaced with a V_{max} (maximum velocity of elimination) of 1,000 mg/hour and a Km (concentration at half maximal velocity) of 250 mg/L. All two-compartment models had an intercompartmental clearance set to 4.0 L/hour. A similar set of models and data sets was previously used to compare NONMEM and Monolix.[17]

ANALYSIS
Software and hardware
Monolix (version 2019R1[14]) using the SAEM algorithm was used as a comparator for the SAEM estimation algorithm implemented in nlmixr. The SAEM algorithm in Monolix was applied using the default settings.

NONMEM (version 7.4.3[9]) using the FOCEI algorithm was used as a comparator for the FOCEI estimation algorithm implemented in nlmixr. The ADVAN13 module was used for estimation with settings of TOL = 6, NSIG = 2, and SIGL = 6 and using the NOABORT and NOOBT options to prevent premature termination of the estimation procedure and to prevent NONMEM's default boundary test for between-subject random effects, respectively. ADVAN13 was used because it implements the Livermore solver for ODEs[18] that is also used in nlmixr. If convergence was not obtained, TOL was increased to 7 or 8. NONMEM was supplemented
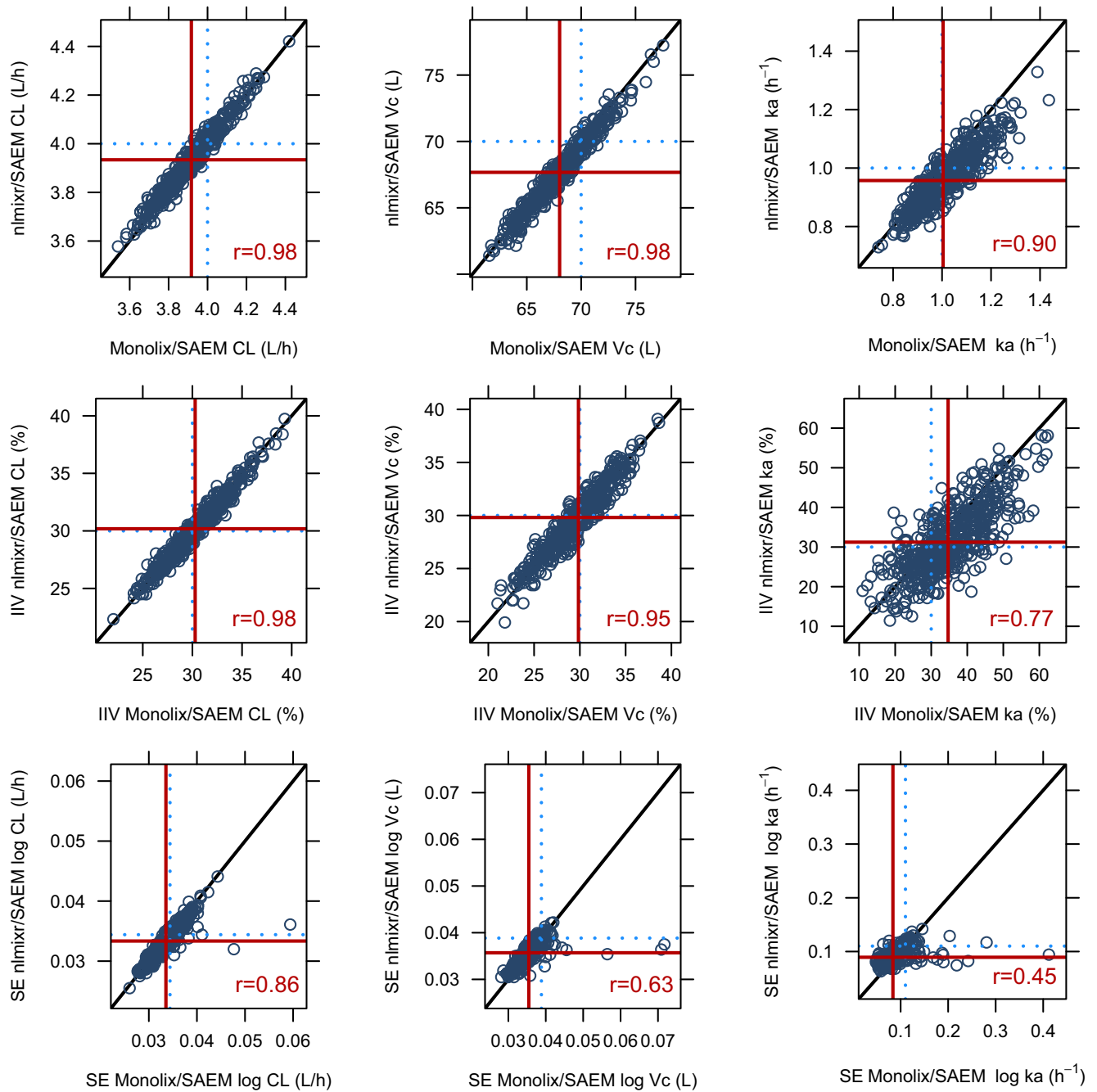
Figure 1 Sparse-data analysis results for nlmixr/stochastic approximation expectation-maximization (SAEM) vs. Monolix/SAEM. Clearance (CL, left column), central volume (Vc, middle column), and absorption rate constant (ka, right column), for population typical parameters (top row), their interindividual variability (IIV, middle row), and the standard error (SE) of their log estimate (bottom row). Dark blue markers: individual paired outcomes for each of the 500 analyses; red lines: median estimated parameter value; blue dotted lines: reference values; black diagonal lines: line of identity.

with PsN, version 4.6.0[15] to sample and analyze the 500 sparse-sample data sets with the bootstrap tool.

R software (64 bit, version 3.6.1[1]) and the RxODE package (version 0.9.1-3 of August 6, 2019) and nlmixr package (version 1.1.1-1 of August 23, 2019[2]) obtained from CRAN were used for parameter estimation. Both the FOCEI algorithm and the SAEM algorithm in nlmixr were applied using the default settings. The 500 sparse sample analyses were

run side by side using the doParallel R package; the required code is available in the **Supplemental Materials**.

All models were implemented using differential equations, and the parameters were estimated on the logarithmic scale. In all cases, IIV was implemented using so-called mu-referencing, in which IIV is expressed linearly with the population typical parameters, and residual error was implemented using the constant coefficient of variation (proportional) residual error model.
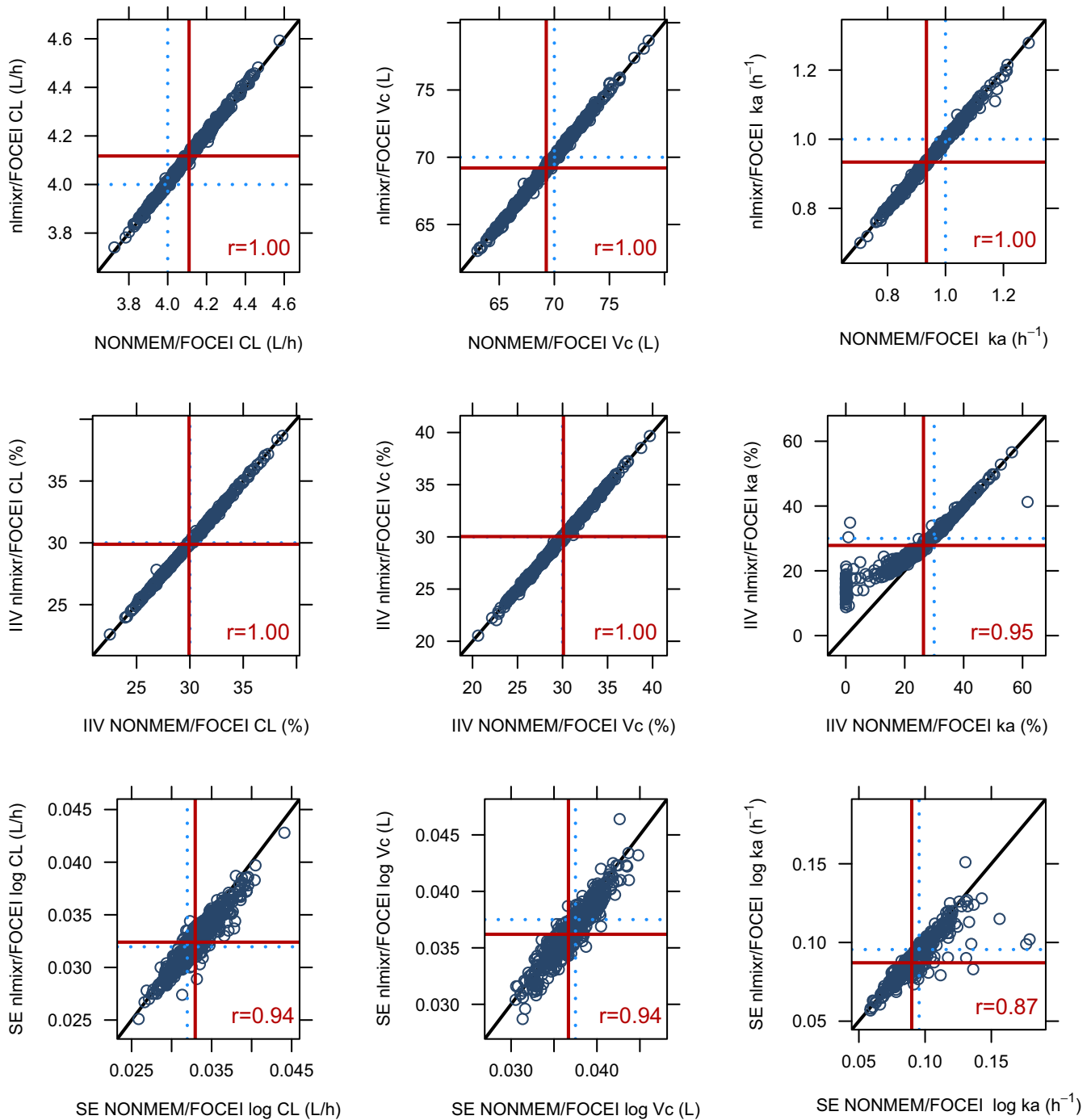
Figure 2 Sparse-data analysis results for nlmixr/first order-conditional estimation with interaction (FOCEI) vs. NONMEM/FOCEI. Clearance (CL, left column), central volume (Vc, middle column), and absorption rate constant (ka, right column) for population typical parameters (top row), their interindividual variability (IIV, middle row), and the standard error (SE) of their log estimate (bottom row). Dark blue markers: individual paired outcomes for each of the 500 analyses; red lines: median estimated parameter value; blue dotted lines: reference values; black diagonal lines: line of identity.

All analyses for all software were run in single-thread mode; parallel computing within a single analysis that is currently only available in NONMEM (as opposed to the side-by-side approach with doParallel as referenced previously) is under active development for nlmixr, but not yet available in the current release.

RESULTS
Sparsely sampled data sets fit using a single model
In the sparsely sampled data scenario with a single one-compartment oral absorption model, fitting the 500 data sets leads to 500 paired outcomes in terms of estimates for population typical parameters, their standard

errors (SEs), and the associated IIV estimates. The correspondence between the results of the two software tools can be visualized using a scatterplot, and a correlation coefficient can quantify the correspondence. By calculating the median estimate across the 500 obtained estimates, a measure of bias can be generated when this median is compared with a reference value. These reference values are readily available for population typical parameters and their IIV because these values were used for simulation. This is not the case for SEs, but a reference for the SEs of the population typical parameters can be obtained by taking the standard deviation of the 500 population typical parameter estimates. **Table S1** in the supplementary materials shows that the SE estimates obtained in this way are very similar between the four estimation algorithms, and consequently, SEs estimated using NONMEM/FOCEI were used as a reference in the figures.

nlmixr/SAEM vs. Monolix/SAEM

The correspondence between estimates obtained using nlmixr/SAEM and Monolix/SAEM is provided in **Figure 1**, with the three rows from top to bottom, population typical parameters ("thetas" in NONMEM parlance), IIVs, and SEs of log-transformed population typical parameters, and the structural parameters (CL, Vc, and ka) in the three columns. The SEs are provided for logarithmic-scaled population typical parameters because estimation takes place on the logarithmic scale.

The results indicate that the estimates of population-typical parameters for the one-compartment sparse-data model are highly correlated, with better correspondence for CL and Vc than for ka. This ordering is the same for IIV and SEs, where the lower correlation coefficient for the SEs are attributable to a small number of outlying SE estimates for Monolix/SAEM.

nlmixr/FOCEI vs. NONMEM/FOCEI

The correspondence between estimates obtained using nlmixr/FOCEI and NONMEM/FOCEI is provided in **Figure 2**. The results indicate that the estimates of population-typical parameters, IIVs, and SEs for the one-compartment sparse-data model are very highly correlated, with the notable exception of IIV for ka (**Figure 2**, middle row, right panel). The data were simulated using an IIV of 30%, and NONMEM/FOCEI resulted in an IIV estimate for ka near zero in 7.6% of the runs, whereas for nlmixr/FOCEI none of the runs estimated an IIV near zero. In practice, near-zero estimates for IIVs are often an indication of insufficient capacity of the software to estimate this parameter. The slightly lower correlation coefficient for SE of ka is attributable to a number of outlying SE estimates for NONMEM/FOCEI.

Figure 3 illustrates that there is a near perfect match in the estimated objective function values between NONMEM/FOCEI and nlmixr/FOCEI. These results suggest that the implementation of FOCEI between the two tools is practically equivalent.

Richly sampled data sets fit using different models

The results from the 36 data-rich models and data set combinations are graphically provided in **Figure 4** using Vc (the only parameter that is common to all 36 models),
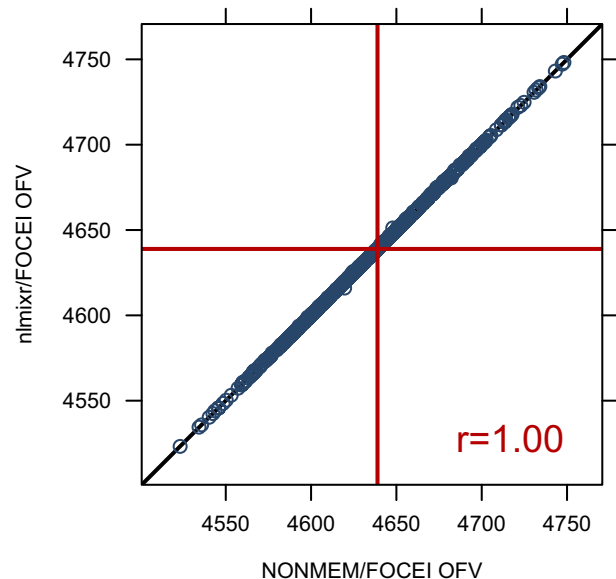

Figure 3 Sparse-data single-model objective function value (OFV) comparison results for nlmixr/first order-conditional estimation with interaction (FOCEI) vs. NONMEM/FOCEI.

where the estimates obtained using Monolix/SAEM, nlmixr/SAEM, NONMEM/FOCEI, and nlmixr/FOCEI are superimposed. The model identifiers are provided on the *x*-axis, and their descriptions are provided in the **Supplemental Materials**; model complexity increases from the left of the graph to the right. This allows a visual assessment if switching from NONMEM/FOCEI or Monolix/SAEM to the corresponding nlmixr algorithm leads to the same results. Graphs for the other pharmacokinetic parameters are provided in the **Supplemental Materials** and show that for population-typical and IIV estimates, the four estimation algorithms provide virtually indistinguishable results.

The SEs are also highly comparable for all four estimation implementations, especially for the less complex models; with increased complexity, both the size of the estimates and the variability increase.

Single-thread run times for the four estimation algorithms are provided in **Figure 5**, where run times are provided on the logarithmic scale to allow comparison with the much longer run times required for the more complex models. Single-threaded run times are similar for Monolix/SAEM when compared with nlmixr/SAEM and longer for NONMEM/FOCEI when compared with nlmixr/FOCEI.

DISCUSSION AND CONCLUSIONS

Extensive comparisons have been performed for the SAEM and FOCEI estimation algorithms implemented within nlmixr with SAEM as implemented in Monolix and FOCEI as implemented in NONMEM, the widely accepted gold standards for nonlinear mixed effects model parameter estimation in pharmacometrics. Scenarios including both sparse sampling for multiple data sets within a single model and rich sampling with a large range of models and inputs were explored.
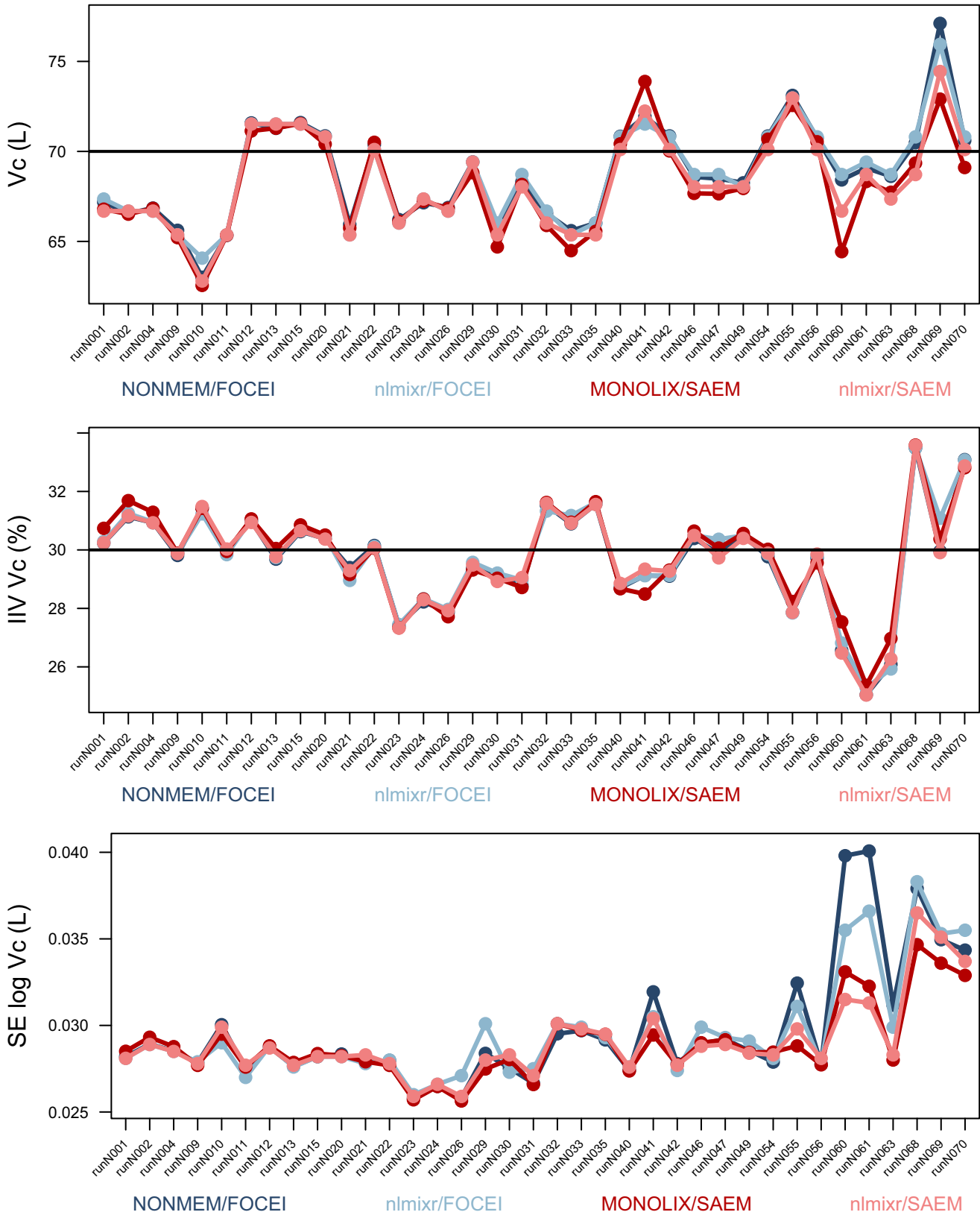
Figure 4 Results for central volume (Vc) for 36 models with richly sampled data sets and multiple models and inputs. Model complexity increases from left to right. Population typical parameters (top), their interindividual variability (IIV, middle), and the standard error (SE) of their log estimate (bottom). Horizontal black line: values used in simulation. FOCEI, first order-conditional estimation with interaction; SAEM, stochastic approximation expectation-maximization.
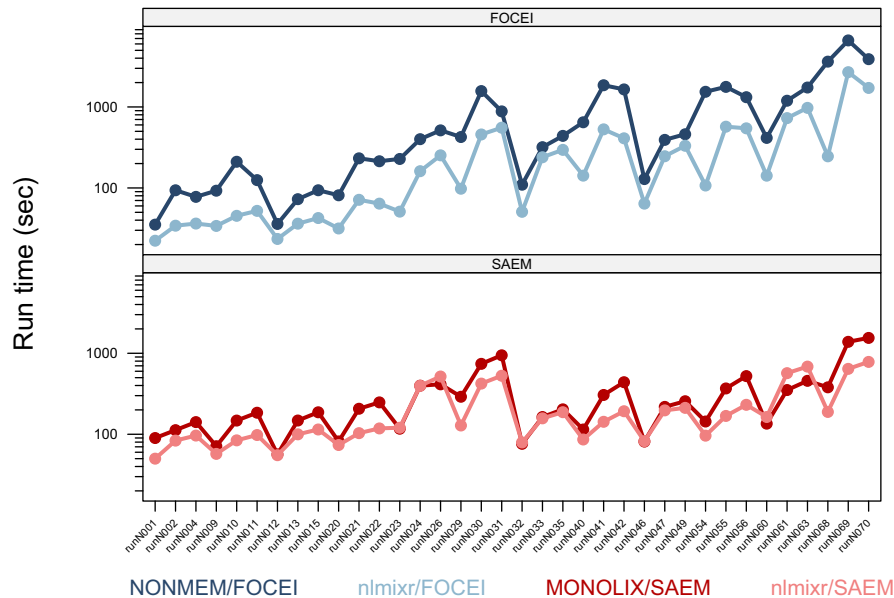
Figure 5 Run times on logarithmic scale using single-thread implementation for the four different estimation algorithms (top row: FOCEI; bottom row: SAEM). FOCEI, first order-conditional estimation with interaction; SAEM, stochastic approximation expectation-maximization.

The results indicate that output is closely comparable across estimation algorithms. Where results deviate, nl-mixr may provide more robust estimates, as witnessed in the absence of outliers for SE estimates and the absence of near-zero estimates for IIV in the sparsely sampled data sets.

The current comparisons were limited to model implementations using ODEs. At the time of writing, solved compartmental systems are available for the nlmixr/nlme and nlmixr/SAEM algorithms, and solved system implementations for nlmixr/FOCEI are under active development. Once implemented for nlmixr/FOCEI, a further increase in computational speed can be expected without impact on the obtained parameter estimates. A parallelized computational implementation for nlmixr/FOCEI and nlmixr/SAEM is also under active development, and preliminary results suggest that the use of multiple computational threads in parallel will deliver a further (substantial) increase in computational efficiency.

These findings provide compelling evidence that nlmixr may be a viable alternative to established offerings for fitting nonlinear mixed effects pharmacometric models. A free and open-source implementation of nonlinear mixed effects modeling algorithms together with robust infrastructure supporting ODE-based model development in a standard R package provides considerable advantages to the pharmacometric community, especially with respect to making state-of-the-art tools and techniques available to researchers and students in resource-limited settings, and may substantially lower the threshold for application of modern statistical and computational techniques in the development of effective medicines.

Supporting Information. Supplementary information accompanies this paper on the *CPT: Pharmacometrics & Systems Pharmacology* website (www.psp-journal.com).

Conflict of Interest. The authors declared no competing interests for this work.

1. R Development Core Team. R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, Vienna, Austria, 2019) <https://www.R-project.org>. Accessed August 24, 2019.
2. Fidler, M. *et al.* nlmixr: nonlinear mixed effects models in population pharmaco-kinetics and pharmacodynamics <https://CRAN.R-project.org/package=nlmixr> (2019). Accessed August 24, 2019.
3. Fidler, M. *et al.* nlmixr: an R package for population PKPD modeling <https://github.com/nlmixrdevelopment/nlmixr> (2019). Accessed August 24, 2019.
4. Wang, W., Hallow, K.M. & James, D.A. A tutorial on RxODE. *CPT Pharmacometrics Syst. Pharmacol.* **5**, 3–10 (2016).
5. Fidler, M. *et al.* RxODE: facilities for simulating from ODE-based models. <https://CRAN.R-project.org/package=RxODE> (2019). Accessed August 24, 2019.
6. Fidler, M. *et al.* nlmixr: nonlinear mixed effects models in population pharmaco-kinetics and pharmacodynamics <https://nlmixrdevelopment.github.io/nlmixr> (2019). Accessed August 24, 2019.
7. Fidler, M. *et al.* Nonlinear mixed-effects model development and simulation using nlmixr and related R open source packages. *CPT Pharmacometrics Syst. Pharmacol.* **8**, 621–633 (2019). https://doi.org/10.1002/psp4.12445.
8. Pinheiro, J., Bates, D., DebRoy, S. & Sarkar, D. nlme: linear and nonlinear mixed effects models <https://CRAN.R-project.org/package=nlme> (2019). Accessed August 24, 2019.
9. Kuhn, E. & Lavielle, M.M. Maximum likelihood estimation in nonlinear mixed effects models. *Comput. Stat. Data Anal.* **49**, 1020–1038 (2005).
10. Beal, S. L., Sheiner, L.B., Boeckmann, A. J. & Bauer, R.J. (Eds). NONMEM 7.4 users guides. (ICON plc, Gaithersburg, MD 1989–2018).
11. Wilkins, J. Graphical diagnostics for pharmacometric models: extension to nlmixr <https://github.com/nlmixrdevelopment/xpose.nlmixr> (2019). Accessed August 24, 2019.
12. Hooijmaijers, R. shinyMixR: shiny dashboard interface for nlmixr <https://github.com/RichardHooijmaijers/shinyMixR> (2019). Accessed August 24, 2019.
13. Hove Rasmussen, C. PharmTeX: an open-source platform for creating PDF reports for the pharmaceutical industry <http://pharmtex.org> (2019). Accessed August 24, 2019.
14. Baltcheva, I. *et al.* ggPMX R package <https://github.com/ggPMXdevelopment/ggPMX> (2019). Accessed August 24, 2019.

15. Monolix version 2019R1 (Lixoft SAS, Antony, France) <http://lixoft.com/products/monolix> (2019). Accessed August 24, 2019.

16. Lindbom, L., Pihlgren, P. & Jonsson, E.N. PsN-Toolkit—a collection of computer intensive statistical methods for non-linear mixed effect modeling using NONMEM. *Comput. Methods Programs Biomed.* **79**, 241–257 (2005)

17. Laveille, C., Lavielle, M., Chatel, K. & Jacqmin, P. Evaluation of the PK and PK-PD libraries of MONOLIX: a comparison with NONMEM. Abstract 1356 <www.page-meeting.org/?abstract=1356> (2008).

18. Hindmarsh, A.C. & Petzold, L.R. LSODA, ordinary differential equation solver for stiff or non-stiff system <http://www.nea.fr/abs/html/uscd1227.html> (2005). Accessed August 24, 2019.