Check for updates

Received: 20 November 2020 | Revised: 6 August 2021 | Accepted: 7 September 2021

DOI: 10.1002/psp4.12714

REVIEW

Metaheuristics for pharmacometrics

Seongho Kim[1] | **Andrew C. Hooker**[2] | **Yu Shi**[3] | **Grace Hyun J. Kim**[3] | **Weng Kee Wong**[3]

[1]Department of Oncology, Wayne State University, Detroit, Michigan, USA

[2]Department of Pharmacy, Uppsala University, Uppsala, Sweden

[3]Department of Biostatistics, University of California Los Angeles, Los Angeles, California, USA

Correspondence
Andrew C. Hooker, Department of Pharmacy, Uppsala University, Box 580, 751 23, Uppsala, Sweden.
Email: andrew.hooker@farmaci.uu.se

Funding information
The research of W.K.W. and Y.S. was partially supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number R01GM107639. S.K. was partially supported by the United States Public Health Service Cancer Center Support Grant P30 CA022453.

Abstract

Metaheuristics is a powerful optimization tool that is increasingly used across disciplines to tackle general purpose optimization problems. Nature-inspired metaheuristic algorithms is a subclass of metaheuristic algorithms and have been shown to be particularly flexible and useful in solving complicated optimization problems in computer science and engineering. A common practice with metaheuristics is to hybridize it with another suitably chosen algorithm for enhanced performance. This paper reviews metaheuristic algorithms and demonstrates some of its utility in tackling pharmacometric problems. Specifically, we provide three applications using one of its most celebrated members, particle swarm optimization (PSO), and show that PSO can effectively estimate parameters in complicated nonlinear mixed-effects models and to gain insights into statistical identifiability issues in a complex compartment model. In the third application, we demonstrate how to hybridize PSO with sparse grid, which is an often-used technique to evaluate high dimensional integrals, to search for D-efficient designs for estimating parameters in nonlinear mixed-effects models with a count outcome. We also show the proposed hybrid algorithm outperforms its competitors when sparse grid is replaced by its competitor, adaptive gaussian quadrature to approximate the integral, or when PSO is replaced by three notable nature-inspired metaheuristic algorithms.

INTRODUCTION

Nature-inspired metaheuristic algorithms have been gaining popularity in the last 2 decades and recently have gained dominant status both in academia and industrial applications.[1–4] Collectively, they now form an important component in artificial intelligence. A main appeal is their simplicity, ease of implementation, and their ability to provide a quality solution to an optimization problem very fast. Another appeal is that these methods often make no assumption on the function to be optimized, with researchers from a widening range of disciplines reporting high success rates of finding an optimal or nearly optimal solution to all types of complex optimization problems in computer science and engineering where there are frequently hundreds or even thousands of variables to optimize.

In the statistical literature, we are also encouraged by recent successes of meta-heuristic algorithms to tackle different types of optimal design problems for different applications. For instance, particle swarm optimization (PSO) or a modified version of PSO was used find various

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
© 2021 The Authors. *CPT: Pharmacometrics & Systems Pharmacology* published by Wiley Periodicals LLC on behalf of American Society for Clinical Pharmacology and Therapeutics.

types of optimal designs for biomedical applications,[5] optimal designs under a nondifferentiable criterion, like a minimax optimal design[6] or a standardized maximin criterion,[7] where the latter involves solving three or four-level nested optimization loops over very different search spaces. PSO-based algorithms were also shown capable of solving discrete optimization problems; for example, modified versions of PSO found (a) optimal supersaturated designs with a large number of factors,[8] (b) optimal repeated measurements over time in a Michaelis-Menten type model,[9] and (c) a class of two-stage adaptive designs that extends the celebrated Simon Two-Stage designs for a phase II trial; the extended designs solved complicated optimization problems involving 10-integer variables subject to multiple nonlinear constraints from the type I error and II error specifications, and allow for none or only one of three alternative hypotheses to be tested at stage 2, depending on the number of responders from stage 1.[10] Other types of metaheuristic algorithms were also recently used to solve optimization problems in statistics, and they include using the Imperialist Competitive Algorithm to find Bayesian optimal designs for a pharmacometric model[11] or a differential evolution (DE) algorithm to find optimal designs in chemometrics.[12] Optimal designs for nonlinear models with many interacting factors in a regression setup were also found using a DE[13] and a variant of PSO, called quantum PSO.[14]

Despite the usefulness and simplicity of metaheuristic algorithms, they seem underutilized in pharmacometrics research. Our main aim in this paper is to demonstrate that metaheuristics is also useful in tackling optimization problems in pharmacometrics, including pharmacokinetics (PKs) and pharmacodynamics (PDs). The intent here is not to give a comprehensive review of the many nature-inspired metaheuristic algorithms in the literature; instead, we focus on one of its members, the widely used PSO. The PSO was proposed by Eberhart and Kennedy[15] and motivated by movement of swarms or flocks of birds. The most basic form of PSO is described in a supplemental file, and as with all such algorithms, there are many modifications after its initial version. Research has shown that PSO has the same effectiveness for finding the true global optimum as the genetic algorithm (GA; another well-known nature-inspired metaheuristic algorithm), but PSO requires significantly fewer function evaluations and, consequently, shorter central processing unit (CPU) time.[16] Another reason for the popularity of PSO is that the default values for the tuning parameters have been shown to work well for a large class of problems,[17,18] including parameter estimation in nonpharmacometrics settings.[19,20] There are comprehensive reviews of PSO and its many variants and their various applications, including for model selection and parameter estimation.[21–25]

Hybridization is a common and often effective procedure to enhance the performance of algorithms by combining two or more of them to tackle a complex optimization problem.[26–28] Optimization problems in pharmacometrics can be quite complex due, in part, to the model nonlinearity, multiple mixed effects, multiple objectives in the optimization, or user-imposed constraints. For example, practical considerations may require that observations from a subject not be taken too close together or that there are varying cost structures from different sites. For such problems, metaheuristics has good potential to give satisfactory solutions where traditional methods fail. For example, hybrid algorithms were proposed for generating optimal designs for discriminating multiple nonlinear models under various error distributional assumptions,[29] and PSO was hybridized with random forest to predict disease progression in patients with idiopathic pulmonary fibrosis.[30]

In what is to follow, we present various applications of PSO that efficiently address optimization issues in pharmacometrics problems. The section Parameter estimation via PSO tackles parameter estimation problems in a complicated nonlinear mixed-effects model. In the section Using PSO to gain insights into statistical identifiability issues, we discuss how PSO can gain insights into statistical identifiability issues in pharmacometrics models with many compartments and parameters. The section Design construction for pharmacometrics models reviews the background for finding efficient designs for pharmacometrics problems and considers a nonlinear mixed-effects model for a count outcome variable, where we wish to find a D-optimal design to estimate all the parameters in the model. Discussion of finding algorithms for optimal designs for such models are limited in the pharmacometrics literature and we propose one here that demonstrates how PSO can be suitably hybridized with sparse grid (SG) for enhanced performance. Because computing the expected information matrix for nonlinear mixed-effects models involved complicated integrals and SG has been shown to be effective for approximating integrals, we hybridize PSO with SG, call the resulting algorithm SGPSO and show it can effectively find efficient designs for mixed effects in nonlinear regression models with count outcomes. Although both PSO and SG have been separately used to find optimal designs, this is the first time they are hybridized for enhanced performance in search for efficient designs in a pharmacometrics setting. We also compare various hybridization schemes and show SGPSO seems to outperform the other schemes for the four-parameter Poisson model. The Discussion section contains a discussion and summarizes the highlights of our review paper.

PARAMETER ESTIMATION VIA PSO

The nonlinear mixed-effect model (NLMEM) is a useful statistical method for analyzing longitudinal data in pharmacometrics. At the first stage of an NLMEM,[31,32] observations are described by a nonlinear subject-specific model:

$$\log y_{ij} = \log f\left(\Phi_i, t_{ij}\right) + \varepsilon_{ij}; \quad i = 1, \ldots, \quad N; j = 1, \ldots, n_i$$

where N is the number of subjects, n_i is the number of observations from the i^{th} subject, y_{ij} the observation (e.g., drug concentration) at time t_{ij}, and $f(\cdot)$ a nonlinear function of a subject-specific parameter vector Φ_i. The residual error term ε_{ij} is an independently and identically distributed normal random variable with mean zero and variance σ^2. Different error structures and distributions are discussed in great detail by Davidian and Giltinan.[33]

At the second-stage, the subject-specific parameter vector Φ_i is modeled as

$$\Phi_i = A_i\beta + B_i b_i$$

where A_i and B_i are known design matrices for a p-dimensional fixed-effects vector β and a q-dimensional random-effects vector b_i, respectively. The random-effects vector is assumed independent and identically distributed as the multivariate normal distribution with q zero-mean variables and variance-covariance matrix Ψ. Alternative distribution assumptions for random effects are discussed in Davidian and Giltinan.[33]

The parameters in the NLMEM described by the above equations are typically estimated by maximum likelihood estimation based on the marginal density of y_i,

$$p(\log y_i|\beta, \sigma^2, \Psi) = \int p(\log y_i|\beta, \sigma^2, b_i)\, p(b_i|\Psi) db_i,$$

where y_i is the vector of observations (e.g., drug concentrations) of i^{th} subject and the conditional density of $\log y_i$ given the random-effects vector b_i is denoted by $p(\log y_i|\beta, \sigma^2, b_i)$, whereas $p(b_i|\Psi)$ is the marginal distribution of b_i.[32] In general, this integral in the marginal density of y_i does not have a closed-form as the function $f(\cdot)$ is nonlinear in b_i so that several different methods are applied for approximating the marginal likelihood. Then, the objective function or likelihood function is numerically minimized with regard to the parameters (β, σ^2, Ψ).

Parameter estimation of NLMEMs is a critical step in the field of pharmacometrics and so in drug development. Several software tools have been developed for estimating pharmacometrics parameters.[32,33-41] Of those, four are widely used in pharmacometrics and they are NONMEM (ICON Development Solutions, Ellicott, MD, USA), Monolix (Lixoft, Orsay, France), Phoenix (Certara, St. Louis, MO, USA), and nlmixr.[42] NONMEM, Monolix, and Phoenix are commercial software tools but Monolix is the only software that provides a free license for noncommercial activities for academic and nonprofit organizations. The software nlmixr is an open-source R package. The estimation approaches used in these tools include, but are not limited to, First Order Conditional Estimation, Stochastic Approximation Expectation-Maximization (SAEM), Monte Carlo Parametric Expectation-Maximization, Importance Sampling Parametric Expectation-Maximization, and Quasi-Random Parametric Expectation-Maximization.[43] Their optimization routines are generally based on expectation-maximization (EM)-like or gradient-based methods. A notable challenge of EM-like or gradient-based methods is getting stuck at saddle points or local optima. Thus, their initial values should be close to the true optimum to achieve a global optimum. Moreover, the EM-like or gradient-based methods are vulnerable to numerical singularities that often occur when the model is nonlinear and/or the parameter domain has many dimensions. These challenges have stimulated the use of metaheuristics for solving global optimization problems using a gradient-free approach in pharmacometrics analysis. Below we discuss the use of four metaheuristic algorithms that have been used in the pharmacometrics literature: Simulated Annealing (SA), GA, Bacterial Foraging Optimization (BFO), and PSO. We briefly review them in the context of the following applications and elaborate further on PSO, which is our focus example of a nature-inspired metaheuristic algorithm in this paper.

For population parameter estimation, a recent application using PSO was suggested by Kim and Li.[31] The strategy, P-NONMEM, is a hybrid approach, using the global metaheuristic search strategy of PSO and the local estimation strategy available in NONMEM. In the developed algorithm, initial values (particles) are generated randomly only for fixed-effects and variance parameters by PSO. NONMEM is then implemented for each particle to find a local optimum for all parameters, including fixed-effects, random-effects, and variance parameters, whereas fixed-effects and variance parameters are guided by PSO for global optimization. They demonstrated that the approach had improved convergence to a global optimum than by using NONMEM or PSO alone. The P-NONMEM algorithm is summarized, briefly, as follows:

- Step I: Initialization

  A swarm of particles of a set of parameters (β, σ^2, and Ψ) is initialized randomly by a multivariate uniform distribution, *Unif*:

$$x_0 = \left(x_0^1, x_0^2, \ldots, x_0^p, \ldots, x_0^P\right) \sim Unif\left(R(\beta) \times R\left(\sigma^2\right) \times R(\Psi)\right), p = 1, \ldots, P,$$

  where P is the size of the initial swarm; $x_0^p = \left(\beta_0^p, (\sigma^2)_0^p, \Psi_0^p\right)$, the p^{th} particle for β, σ^2, and Ψ; $R(z)$ is the range of a random variable (vector) z.

- Step II: NONMEM estimation

  For particle p at the k^{th} iteration, a NONMEM-based local estimation is performed. Let the current position of particle p at iteration k, $x_0^p = \left(\beta_k^p, (\sigma^2)_k^p, \Psi_k^p\right)$, be an initial value for the NONMEM estimation. Their estimates $\hat{x}_k^p$ are obtained along with the random-effect estimate $\hat{b}_k^p$ through NONMEM. Then the current position x_k^p is updated with the convergence estimate $\hat{x}_k^p$, i.e., $x_k^p \leftarrow \hat{x}_k^p$.

- Step III: Local and global best positions

  The goodness-of-fit (GOF) for the p^{th} particle at the k^{th} iteration (GOF_k^p) is calculated given the update current position x_k^p and the random effect estimate $\hat{b}_k^p$, $\text{GOF}_k^p = -2\log L(y|x_k^p, \hat{b}_k^p)$. GOF_k^p is then compared to the best previous local and global best goodness of fits (i.e., $\text{GOF}_{\text{lbest}}^p$ and $\text{GOF}_{\text{gbest}}$), and the current local and global bests are updated as follows:

  - Updating the local best position: $x_{\text{lbest}}^p \leftarrow x_k^p$ and $\text{GOF}_{\text{lbest}}^p \leftarrow \text{GOF}_k^p$ if $\text{GOF}_k^p < \text{GOF}_{\text{lbest}}^p$.
  - Updating the global best position $x_{\text{gbest}} \leftarrow x_k^p$ and $\text{GOF}_{\text{gbest}} \leftarrow \text{GOF}_k^p$ if $\text{GOF}_k^p < \text{GOF}_{\text{gbest}}$.

- Step IV: Convergence

  If iteration k reaches to the user-specified maximum K or all the particles converge to its global best, P-NONMEM will stop. Otherwise, it will go to step V.

  - Stop if $k = K$ or $\left|x_{\text{gbest}} - x_k^p\right| < \varepsilon$, for all $\varepsilon > 0$ and for all $p = 1, \ldots, P$
  - Go to step V otherwise

The current positions $\{x_k^p\}_{p=1}^P$ will be updated to $\{x_{k+1}^p\}_{p=1}^P$ by PSO until it converges. That is, $x_{k+1}^p \leftarrow x_k^p + v_{k+1}^p$, where v_{k+1}^p is the updated velocity in PSO.

For parameter estimation in the nonpopulation setting a number of different hybrid metaheuristic algorithms have been developed. Türksen and Tez[44] developed two hybrid algorithms, which are called GANMS and PSONMS, to estimate the parameters of compartment models. GANMS and PSONMS are combination of a derivative-free local optimization, Nelder-Mead Simplex (NMS),[45] and GA and PSO, respectively. The local search of GA and PSO was replaced with NMS so that both GA and PSO contribute to the global search only. They further compared the performance in terms of parameter estimation and suggested that PSONMS performs better than GA, PSO, and GANMS. They also used GA to develop an approach for bias-corrected point estimates and bias-corrected accelerated confidence interval estimates of two-compartment models.[46] To do this, they used a bootstrap method upon point estimates obtained by a derivative-based nonlinear least squares (NLS) approach or GA. Their comparison showed that the parameter estimates obtained by GA are quite unbiased compared to those obtained by NLS. Recently, Pan et al.[47] developed a support vector regression model with PSO (PSO-SVR) to simultaneously characterize the PKs and PDs of Moutan Cortex and Moutan Cortex charcoal widely used in traditional Chinese medicine. In PSO-SVR, PSO was used to find parameters that fit the chemical component to the drug efficacy by optimizing the support vector regression model. They further compared PSO-SVR with a back-propagation neural network, which is one of the artificial neural networks and has better flexibility and accuracy, demonstrating that PSO-SVR is a better approach for pharmacometrics than the back-propagation neural network. Toney[48] used SA, which is implemented in a biochemical system simulator COPASI (http://www.copasi.org), to determine free energy profiles for alanine racemase and triosephosphate isomerase with a complex kinetic mechanism (such as Michaelis-Menten kinetics). Sowparnika et al.[49] used bacterial foraging-oriented PSO (BFOA-PSO) to develop an automatic drug infusion control system during cardiovascular surgery. BFOA-PSO[50] is a hybrid of BFO[51] for determining a solution by elimination and dispersal stage, and PSO for social information trade-off. Ye et al.[52] utilized PSO to estimate parameters of a three-compartment toxicokinetic model for the absorption, distribution, metabolism, and elimination of tricothecenes-2 toxin (T-2) in shrimp.

USING PSO TO GAIN INSIGHTS INTO STATISTICAL IDENTIFIABILITY ISSUES

Parameter estimation has typically two types of identifiability issues; mathematical and statistical identifiability. Mathematical identifiability, also called structural or deterministic identifiability, is the ability to identify model

parameters in noise-free data free from design constraints (unlimited data). Statistical identifiability, also known as numerical identifiability, is the identifiability of parameters estimated from noisy data, with design constraints.[53–55] Mathematical identifiability is largely not a challenging problem in parameter estimation in pharmacometrics, as pharmacometrics models are developed using general differential equations with solid theoretical and mathematical foundations. The critical problem with identifiability of pharmacometrics modeling mainly stems from statistical identifiability due to lack of samples and high residual variability.

For example, in pharmacometrics studies, the Michaelis-Menten (MM) equation is often used to describe intrinsic clearance:

$$\mathrm{CL}_{\mathrm{int}} = V_{max} / \left(Km + C\left(t\right)\right),$$

where V_{max} is the maximum enzyme activity; Km an inverse function of the affinity between drug and enzyme; $C\left(t\right)$ an unbound drug concentration. Km is also called the MM constant having the units of $C\left(t\right)$. However, when Km is much larger than the concentration $C\left(t\right)$ (i.e., $Km \gg C\left(t\right)$), $\mathrm{CL}_{\mathrm{int}} \cdot C\left(t\right)$ becomes approximately equal to $\left(V_{max}/Km\right) \cdot C\left(t\right)$. In addition, when Km is much smaller than the concentration $C\left(t\right)$ (i.e., $Km \ll C\left(t\right)$), $\mathrm{CL}_{\mathrm{int}} \cdot C\left(t\right)$ is approximately equal to $V_{max} \cdot C\left(t\right)$. In other words, if the concentration $C\left(t\right)$ is either much less or greater than Km, one will not be able to estimate both Km and V_{max} separately due to statistical identifiability.[54,55]

Numerous methods have been presented to deterministically detect unidentifiable parameters, such as Laplace transforms,[56] similarity transform approaches,[57] Voterra and generative power series approaches,[58] differential algebra approaches,[59] and alternating conditional expectation algorithms,[60] but less development has been made for statistical identifiability.[61–63] One practical approach to evaluating statistical identifiability is local sensitivity analysis. It uses the first partial derivative of the differential equation with respect to the parameters, and depends on the nonsingularity of the Fisher information matrix corresponding to the Taylor series method and the differential algebra method.[64,65] However, if the estimate is far from the actual value or the model has very complex dynamics, local sensitivity analysis is likely to make a wrong decision. Thus, a global sensitivity analysis for robust experimental design was proposed by Yue et al.[66] based on the modified Morris method. Although Bayesian approaches do not have mathematical or statistical identification problems thanks to priors, statistical identifiability can lead to poor convergence on parameter estimation.[67,68] To speed up the convergence rate of Bayesian approaches, the Switching Monte Carlo Markov Chain (MCMC) method was introduced by Kim and Li[54] to estimate parameters regardless of the statistically identifiable situations. It is a Bayesian approach and to randomly switch between two schemes, single component MCMC and group component MCMC. However, all of the aforementioned approaches still require initial guesses or prior knowledge regarding the underlying relationship of the parameters.

For this reason, Kim and Li[55] recently introduced an optimization approach using PSO that not only handles identifiability as a whole but also does not require preprocessing to obtain initial guesses or prior knowledge. Their developed algorithm, called LPSO, is a combination of PSO and a derivative-free local optimization, which is a simplex search algorithm,[45] to enhance the convergence rate of the local best. Consequently, LPSO converges to a global optimum much faster than PSO and can be applied to parameter estimation regardless of statistical or numerical singularity. They further introduced several convergence diagnostic measures to detect when to stop LPSO and to indicate whether the pharmacometrics model is statistically identifiable or not.

LPSO was, in fact, used to aid in the estimation of parameters in a complex PK model.[69] In this work, the objective was to integrate dynamic positron emission tomography and conventional plasma PK studies to characterize the plasma and tissue PKs of 1-(2-deoxy-2-fluoro-beta-d-arabinofuranosyl) uracil (FAU) and 1-(2-deoxy-2-fl uoro-beta-d-arabinofuranosyl) 5-methyluracil monophosphate (FMAU). An 8-compartment model was used to simultaneously characterize the plasma and tissue pharmacokinetics of FAU and its active metabolite FMAU. The parameter estimation was performed mainly using SAEM in combination with MCMC, as implemented in Monolix (Lixoft, Orsay, France). In the study there were 12 subjects and 20 parameters to be estimated: V_1, V_2, V_3, V_4, V_5, V_6, V_7, V_8, CL_{12}, CL_{13}, CL_{14}, CL_{56}, CL_{57}, CL_{58}, $Vmax_1$, $Vmax_2$, $Vmax_3$, CL_1, CL_2, and Km, where V indicates a volume of a compartment; CL a clearance; V_{max} a maximum velocity; and Km an MM constant. There are six differential equations that include the MM equation which can cause issues with statistical identifiability. The complexity of the PK model and the lack of literature-based historical data made it difficult for SAEM to estimate PK parameters in the NLMEM. In particular, SAEM-based parameter estimation became very difficult because the statistical identifiability of each PK parameter was unknown and reasonable initial values and/or boundaries were not available. To resolve these difficulties, LPSO was applied to not only find good initial values and/or boundaries for each PK parameter but also to enable the researcher to investigate statistical identification problems for the MM equation in a more in-depth manner.

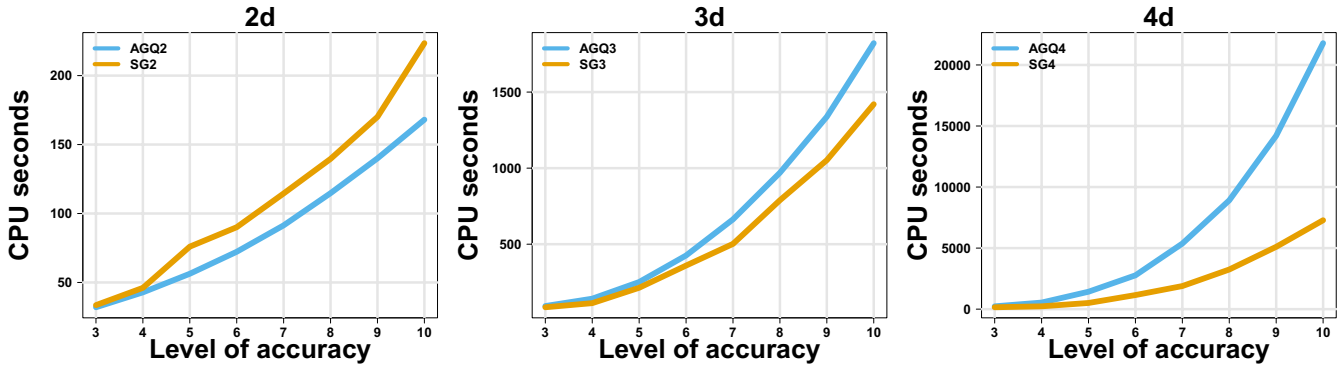

FIGURE 1 CPU times (seconds) required by the AGQ and SG for evaluating two to four parameter Poisson-type models in the section Nonlinear mixed-effects Poisson-type models at different accuracy levels. AGQ, adaptive Gaussian quadrature; CPU, central processing unit; SG, sparse grid

DESIGN CONSTRUCTION FOR PHARMACOMETRICS MODELS

Background

NLMEMs are increasingly used to understand drug effects and finding an efficient design at minimum cost can be challenging. If the goal is to estimate parameters in a model, *D*-optimality is appropriate and a *D*-optimal design provides the most accurate estimates of the parameters among all designs using minimal resources. There are numerical approaches for finding optimal designs, but for complicated models, such as NLMEMs, many current algorithms can be slow or unable to find a *D*-efficient design. In pharmacometrics, the problem can be especially computationally challenging for models with a discrete end point.

At the design stage, the NLMEMs in pharmacometric experiments contain unknown parameters and so the objective function cannot be optimized directly. A simple method to overcome the problem is to replace the unknown parameters by their nominal parameter values from previous experiments on similar drugs or from a pilot study. The objective function can then be optimized and the resulting designs are called locally optimal because they depend on the nominal values.[70] Typically, the design criterion is formulated in terms of the expected Fisher information matrix (FIM), which measures the worth of the design.[71–74] For NLMEMs, the FIM does not have an analytic solution so that the integral has to be approximated before the criterion is optimized.

A common way to approximate the FIM is to use a first order (FO) linearization of the model around the expectation of random effects.[75,76] This has been shown to work well and the method is available in several software programs.[77] Although this method is generally efficient,[78–80] FO has limitations when it is applied to NLMEM with discrete outcomes.[81] Recent work has suggested that adaptive Gaussian quadrature (AGQ) can be useful for such models with a small number of random effects, but the calculation of the FIM can become slow, especially when the number of random effects in the model increases (the curse of dimensionality).[82]

The SG techniques for numerical integration has been shown to be efficient for evaluating high dimension integrals and is less effective for low dimensional integrals. Of particular relevance is the work of Plumlee,[83] who provided an excellent background with technical details of SG with compelling examples and asserts that "The computational savings can be several orders of magnitude when the input is located in a high-dimensional space." See also introductory information and illustrative applications of SG at https://www.ucl.ac.uk/uctpjyy/downloads/SparseGrid.pdf, where instructions on how to implement and use the SG codes from CRAN are available. Our findings here support the comparative advantages of SG for computing the FIM when we want to approximate a high dimensional integral (Figure 1).

Following convention, the population log-likelihood is dependent on data derived from the population design, which is made up of a collection of elementary designs, where each may have multiple individuals with the same design. As an example, a population design has E elementary designs $\xi_e, e = 1, 2, \ldots, E$. If there are N_e subjects assigned to $\xi_e, e = 1, 2, \ldots, E$, then the population design can be written as $\Xi = \{[\xi_1, N_1];[\xi_2, N_2];\ldots;[\xi_E, N_E]\}$ with the constraint that $\sum_{e=1}^{E} N_e = N$. If all subjects have identical elementary designs, then $E = 1$ and the population design is simply $\Xi = \{\xi, N\}$. Assuming independence across the N subjects and Ψ^* is the unknown model parameters in the NLMEM, the population FIM, $M_{\mathrm{PF}}\left(\Psi^*, \Xi\right)$, for a design Ξ, is the sum of N elementary matrices $M_F\left(\Psi^*, \xi_i\right)$ from each subject, that is:

$$M_{\mathrm{PF}}\left(\Psi^{*},\Xi\right)=\sum_{i=1}^{N}M_{F}\left(\Psi^{*},\xi_{i}\right), \tag{1}$$

where the elementary matrices are the expectation across all possible datasets E_{y_i} of the second derivative of the log-likelihood surface, or equivalently

$$M_{F}\left(\Psi^{*},\xi_{i}\right)=E_{y_i}\left(\frac{\partial \log L\left(\Psi^{*},y_{i}\right)}{\partial\Psi}\frac{\partial \log L\left(\Psi^{*},y_{i}\right)}{\partial\Psi^{T}}\right). \tag{2}$$

Our criterion is *D*-optimality, which is what we want to find as a design that maximizes the determinant of the FIM after it is normalized by the number of parameters to be estimated. When nominal values Ψ_0 for the model parameters are available, we replace Ψ^* in the FIM by Ψ_0 and the FIM for the design ξ becomes $M_{\mathrm{PF}}\left(\Psi_0,\Xi\right)$. We then optimize the following:

$$\Xi_{D}=\underset{\Xi\in\Xi^{*}}{arg\,max}\left\{\left|M_{\mathrm{PF}}\left(\Psi_{0},\Xi\right)\right|^{1/p}\right\}, \tag{3}$$

where |.| denotes the determinant, Ξ^* is the set of all designs on the design space, and p is the number of estimated parameters in the model. This normalized FIM determinant is the *D*-criterion value; larger values indicate better designs. To compare the performance of two designs Ξ_1 and Ξ_2 for estimating the model parameters, we use the ratio:

$$\left\{\frac{\left|M_{\mathrm{PF}}\left(\Psi_{0},\Xi_{1}\right)\right|}{\left|M_{\mathrm{PF}}\left(\Psi_{0},\Xi_{2}\right)\right|}\right\}^{1/p}, \tag{4}$$

which is the *D*-efficiency of the design Ξ_1 relative to the design Ξ_2. As an example, if the above ratio is 0.5 or 50%, this means that the design Ξ_1 needs twice as many observations for it to do as well as the design Ξ_2.

In the population design setting, where the outcome is usually continuous, traditional optimization methods have been used to search for efficient designs.[84] Hybridized heuristic methods for finding efficient designs that tend to work well on a wider set of problems have also been used. For more complex design problems, multiple algorithms have been applied in sequence.[85,86] However, apart from Ueckert,[82] who used an MCMC approach, algorithms for finding efficient designs for NLMEMs with a count end point seem to be less discussed in the design setting. We next present a few specific models with a count end point and show that a hybridization of SG and PSO, called SGPSO, can find *D*-efficient designs for these types of models.

Nonlinear mixed-effects Poisson-type models

Inspired by a previous work,[82] we consider Poisson NLMEM with various dose levels as an example system. The responses recorded over time are non-negative integer values (counts, k). The simplest two-parameter model investigated describes a response probability at dose level $D=D_i$ by:

$$p(y_i=k|b_i,\Psi)=\frac{\lambda_i\exp\left(\lambda_i\right)}{k!}\text{ with }\quad\log\left(\lambda_i\right)=\theta_1\left(1-\frac{D_i}{D_i+\theta_2}\right)$$

where θ_1, θ_2 follow a log-normal distribution: $(\theta_1,\theta_2)^T=(\mu_1,\mu_2)^T\exp\left((b_1,b_2)^T\right)$. We assume $(\mu_1,\mu_2)^T=(1,0.5)$ and $(b_1,b_2)^T\sim N(\mathbf{0},\ \mathrm{diag}(0.09,0.09))$.

A more complicated version of this model is a 3-parameter model:

$$p(y_i=k|b_i,\Psi)=\frac{\lambda_i\exp\left(\lambda_i\right)}{k!}\text{ with }\quad\log\left(\lambda_i\right)=\theta_1\left(1-\frac{\theta_3 D_i}{D_i+\theta_2}\right)$$

where θ_1, θ_2, θ_3 follow a log-normal distribution: $(\theta_1,\theta_2,\theta_3)^T=(\mu_1,\mu_2,\mu_3)^T\exp\left((b_1,b_2,b_3)^T\right)$. We assume $(\mu_1,\mu_2,\mu_3)^T=(1,\ 0.5,\ 0.1)$ and $(b_1,b_2,b_3)^T\sim N(\mathbf{0},\ \mathrm{diag}(0.09,0.09,0.01))$.

We also consider a four-parameter model:

$$p(y_i=k|b_i,\Psi)=\frac{\lambda_i\exp\left(\lambda_i\right)}{k!}\text{ with }\quad\log\left(\lambda_i\right)=\theta_1\left(1-\theta_4-\frac{\theta_3 D_i}{D_i+\theta_2}\right)$$

where θ_1, θ_2, θ_3 follow a log-normal distribution: $(\theta_1,\theta_2,\theta_3)^T=(\mu_1,\mu_2,\mu_3)^T\exp\left((b_1,b_2,b_3)^T\right)$ and θ_4 follows a normal distribution. Further, let $(\mu_1,\mu_2,\mu_3)^T=(1,\ 0.5,\ 0.1)$, $(b_1,b_2,b_3)^T\sim N\left(\mathbf{0},\mathrm{diag}\left(0.09,\ 0.09,\ 0.01\right)\right)$ and $\theta_4=\mu_4+b_4$ where $\mu_4=0$ and $b_4\sim N\left(0,0.01\right)$.

For each model, we assume designs with 20 subjects and 90 observations per subject at different dose levels between 0 and 1.[82] The reference design ξ_1 provides 30 observations at each of the three doses at 0, 0.4, and 0.7. In terms of notation for a population design, we have $E=1, N=20$ with one elementary design (i.e., $\Xi=\left\{\xi_1,N\right\}$).

A comparison of using AGQ and SG for Poisson NLMEM

We compare the performance of AGQ and SG for approximating FIMs of Poisson NLMEM using models with discrete outcomes and different numbers of parameters and varying number of random effects. Following Ueckert et al.,[82] we use quasi random Monte Carlo (QRMC) to integrate out the random effects. All other conditions

were fixed between SG and AGQ, including using 1000 QRMC samples and the same one-dimensional three-point Gaussian quadrature. The comparison criteria are CPU times required to compute the FIM and the different numbers of node points required at different accuracy levels. An SG constructed with accuracy level k integrates complete polynomials of total order $2k-1$ exactly. SG limits the total order of all the one-dimensional monomials used to approximate the multidimensional integral using one-dimension polynomials of lower degrees, so they are less accurate, but require fewer nodes to approximate when the integral is high-dimensional. For example, if the integral is nine-dimensional, SG requires 5965 nodes and the product rule for AGQ requires 195,3125 nodes. Consequently, the choice of k is a compromise between computationally efficiency (having fewer nodes to evaluate the integral) and approximation accuracy. A more complete and technical definition of accuracy level is given in Shi.[87]

Figure 1 shows the CPU times required by AGQ and SG to compute the FIM at different accuracy levels for the two to four parameter Poisson-type models. Regardless of the accuracy levels, AGQ outperforms SG in the two-dimensional model, but SG outperforms AGQ for the three-dimensional model, and much more so for the four-dimensional model; at the accuracy level 10 for the four-dimensional model, SG requires a third of the time required by AGQ.

Figure 2 shows the criterion values for the generated designs for the two to four parameter models at different accuracy levels from the AGQ- and SG-generated designs. They are the determinants of the normalized FIMs and their values get closer when the accuracy level increases. At higher accuracy levels, the difference becomes indistinguishable. When the model has more parameters, higher accuracy levels are required for two criterion values to be close. The figure also shows AGQ outperforms SG in terms of criterion values consistency across accuracy levels and dimension of the models. Thus, the SG method should be seen as a way to decrease computational time, and care should be taken with calculations at low accuracy levels.

Algorithms with different hybridizations for finding D-efficient designs for the four-parameter Poisson NLMEM

This subsection compares results using different hybridization schemes for searching for a D-efficient design for the four-parameter Poisson NLMEM in the previous subsection. The FIM is approximated by either QRMC-SG (SG-PSO) or QRMC-AGQ (AGQ-PSO). We used 1000 QRMC samples and used the same one-dimensional three-points Gaussian quadrature for AGQ and SG. The accuracy level used for SG in this example is four. PSO was then used to optimize the dose levels and observation allocations based on D-optimality for two scenarios. The first is a "Fix allocation" scenario, where we fix the observation allocation to be the same as the reference design, and find the three dose levels that maximize the D-criterion value. The second is the "Flexible" scenario, where we search for the optimal allocation scheme without restriction. The PSO parameters used for this example were as follows: 100 maximum number of iterations, $c_1 = c_2 = 2$, the inertia parameter was set to linearly decrease from 0.9 to 0.4 and the swarm size was 40 particles.[88]

Table 1 reports the results for the two design scenarios. We observe that all designs are more efficient than the reference design with relative D-efficiencies ranging from 131% to 135%. The SG-PSO approach finds relatively equivalent designs compared to AGQ-PSO, but in terms of CPU time, AGQ-PSO required 1.7 times the CPU time compared to SG-PSO.

We next compare when SG was used but replaced PSO by three of its strong competitors: GA, DE and a version of the quantum-inspired PSO (QPSO). We used these

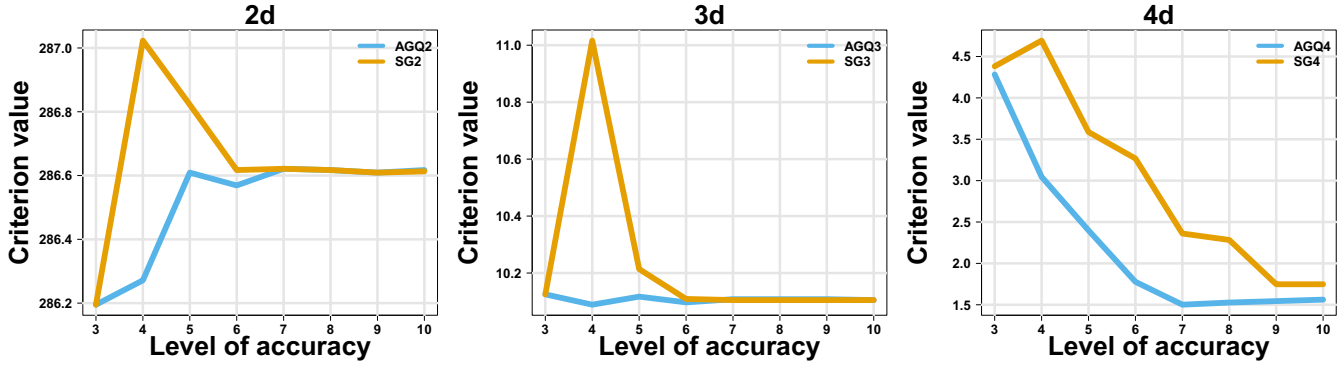

FIGURE 2 Criterion values of the generated designs for the two to four parameter models in the section Nonlinear mixed-effects Poisson-type models at different accuracy levels

TABLE 1 D-optimality criterion values of the AGQ-PSO- and SG-PSO-generated designs (at accuracy level 4) for the four-parameter Poisson-type model, their D-efficiencies relative to the reference design in parentheses (relative efficiency, RE) and CPU time in seconds to find the generated designs

	Dose levels	Number of observations	D-criterion (RE)	CPU time
Reference design	0	30	57.691	N/A
	0.4	30		
	0.7	30		
Fix allocation – AGQ-PSO	0	30	73.098 (126.7%)	27.164.47
	0.25	30		
	0.88	30		
Fix allocation – SG-PSO	0	30	75.694 (131.2%)	16.203.58
	0.25	30		
	0.94	30		
Flexible – AGQ-PSO	0	19	76.699 (132.9%)	27.455.37
	0.21	29		
	0.78	42		
Fix – SG-PSO	0	19	76.699 (132.9%)	16.259.36
	0.21	29		
	0.78	42		

Note: The fixed allocation scheme designs require the algorithm to find the three best dose levels with an equal number of observations and the flexible scheme allows the algorithm to determine the three best dose levels and the optimal number of observations at each dose, given that in both cases, the total number of observations per individual is 90.

Abbreviations: AGQ, adaptive Gaussian quadrature; CPU, central processing unit; N/A, not applicable; PSO, particle swarm optimization; RE, relative efficiency; SG, sparse grid.

algorithms coupled with SG to find efficient designs for the same four-parameter Poisson NLMEM. We consider the "Flexible" scenario. The parameters of GA are set as follows: the population size is 50, the number of generations is 100, elitism is 2, cross-over probability is 0.8, and mutation probability is 0.1. The parameters of DE are set as follows: the population size is 80, the number of generations is 100, step size is 0.8, and cross-over probability is 0.5. The parameters of QPSO are set as follows: the alpha is linearly decreasing from 1.4 to 0.4, the population size is 40, the maximum number of iterations is 100. These are generally default parameters for the algorithms.

Table 2 displays fixed allocation scheme designs found when SG is hybridized with PSO and three other metaheuristic algorithms: QPSO, GA, and DE. It is clear that SG-PSO finds a design with the highest *D*-criterion value. Note that these results are clearly dependent on the algorithm parameters and should be seen as a guide rather than a full comparison.

TABLE 2 A comparison of algorithms when SG, with accuracy level four, is hybridized with other metaheuristic algorithms for finding efficient designs for the four-parameter Poisson-type model

	Dose levels	Number of observations	D-criterion (relative efficiency)
PSO	0	19	76.70 (100%)
	0.21	29	
	0.78	42	
QPSO	0	46	65.08 (84.9%)
	0.26	29	
	0.97	15	
GA	0.02	29	69.89 (91.1%)
	0.31	30	
	0.88	31	
DE	0.01	7	70.24 (91.6%)
	0.22	47	
	0.97	36	

Note: Each algorithm finds the best three dose levels and the optimal number of observations at each dose, subject to the constraint that they sum to 90.

Abbreviations: DE, differential evolution; GA, genetic algorithm; PSO, particle swarm optimization; QPSO, quantum-inspired particle swarm optimization; SG, sparse grid.

DISCUSSION

We have demonstrated the utility of nature-inspired metaheuristic algorithms for tackling a few optimization problems in pharmacometrics. For example, finding optimal designs for NLMEMs is a notoriously difficult problem; the FIM has to be numerically approximated before the design is found by optimizing a function of the FIM. There is no theory to confirm optimality for such designs and a common practice is to implement the design with the

best criterion value among a few recommended designs by medical experts in the area. Metaheuristics is likely able to find an efficient design for such models, or a properly hybridized version of it can do so. We demonstrated an application by combining SG and PSO and showed the hybridized version can find a *D*-efficient design for a Poisson model with multiple parameters. We also combine SG with a few other metaheuristic algorithms, such as QPSO, GA, and DE, and find that the SG-PSO hybridization appears to outperform these other methods. Our examples show that SG-PSO-generated designs outperformed the reference designs in all instances in terms of *D*-efficiencies. These results also match previous findings for different models with a continuous outcome.[87]

We note that designs found by PSO or SG-PSO should be referred to as PSO-generated designs or SG-PSO-generated designs, and not *D*-optimal designs. This is because there is no guarantee that a metaheuristic algorithm will find an optimal design. Under such circumstances when the true optimum is unknown, we recommend using several types of metaheuristic algorithms and hope that they all generate similar designs. Otherwise, we can hybridize a metaheuristic algorithm with another optimization algorithm so that the hybridized version works better than either one of them alone. This optimization process is implemented in software for population optimal designs (PopED,[86] https://andrewhooker.github.io/PopED/).

We conclude by noting that metaheuristics should be used as a last resort when traditional optimization methods fail. In complex situations, the optimization problem may not be able to be formulated into one that meets all assumptions required by the traditional methods. It is in these situations that metaheuristics should be used. However, because metaheuristic algorithms are usually fast, flexible, easy to implement, and use, many researchers have used them for convenience. In addition, we note that not all metaheuristics perform equally well for all problems and some are harder to use because, in part, they are motivated differently or have different numbers of tuning parameters. Tuning the parameters for a metaheuristic algorithm for convergence can be both difficult and very time consuming, and is a perennial problem for metaheuristics. Similarly, trying to rigorously prove a nature-inspired metaheuristic algorithm converges to the global optimum has been a stubbornly difficult problem. Both these issues are active research areas in metaheuristics, even as incremental advances are constantly being made.[89,90] In practice, however, one can frequently and convincingly argue that only an approximate optimal solution is needed. Other open problems in metaheuristics with particular relevance for health care studies are described in Tsai.[91]

CONFLICT OF INTEREST

The authors declared no competing interests for this work.

REFERENCES

1. Yang XS. *Nature-inspired Metaheuristic Algorithms*. Luniver Press, 2010.
2. Whitacre JM. Recent trends indicate rapid growth of nature-inspired optimization in academia and industry. *Computing*. 2011;93:121-133.
3. Whitacre JM. Survival of the flexible: explaining the recent dominance of nature-inspired optimization within a rapidly evolving world. *Computing*. 2011;93:135-146.
4. Abdel-Basset M, Sangaiah AK. Chapter 10 – Metaheuristic algorithms: a comprehensive review. In: Kumar SA, Michael S, Zhiyong Z, eds. *In Intelligent Data-centric Systems. Computational Intelligence for Multimedia Big Data on the Cloud with Engineering Applications*. Academic Press; 2018:185-231.
5. Jiaheng Q, Ray-Bing C, Weichung W, Wong WK. Using animal instincts to design efficient biomedical studies. *Swarm Evolut Computat J*. 2014;18:1-10.
6. Chen R-B, Chang S-P, Wang W, Tung H-C, Wong WK. Minimax optimal designs via particle swarm optimization methods. *Stat Comput*. 2015;25(5):975-988.
7. Chen RB, Chen PY, Wong WK. Standardized maximin D-optimal designs for pharmacological models via particle swarm optimization techniques. *Chemometr Intell Lab Syst*. 2018;169:79-86.
8. Phoa FKH, Chen R-B, Wang W, Wong WK. Optimizing two-level supersaturated designs using swarm intelligence techniques. *Technometrics*. 2016;58(1):43-49.
9. Bing CR, Yang CP, Heng-Chin T, Kee WW. Exact D-optimal designs for Michaelis-Menten model with correlated observations by particle swarm optimization. *Festschrift in Honor of Hans Nyquist on the Occasion of His 65th Birthday*. 2015;60.
10. S, Wong WK. Extended two-stage adaptive designs with three target responses for phase II clinical trials. *Stat Methods Med Res*. 2018;27(12):3628-3642.
11. Masoudi E, Holling H, Duarte BPM, Wong WK. Metaheuristic adaptive cubature based algorithm to find Bayesian optimal designs for nonlinear models. *J Computat Graph Stat*. 2019;28(4):861-876.
12. Stokes Z, Mandal A, Wong WK. Using differential evolution to design optimal experiments. *Chem Intell Lab Syst*. 2020;199:103955.
13. Xu WK, Tan KC, Xu JX. Finding high-dimensional D-optimal designs for logistic models via differential evolution. *IEEE Access*. 2019;7(1):7133-7146.
14. Lukemire J, Mandal A, Wong WK. Optimal experimental designs for ordinal models with mixed factors for industrial and healthcare applications [published online ahead of print November 2, 2020]. *J Quality Technol*. 10.1080/00224065.2020.1829215
15. Russell E, James K. A new optimizer using particle swarm theory. MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science 1995; 39-43.
16. Rania H, Babak C, Olivier W, Gerhard V. A comparison of particle swarm optimization and the genetic algorithm. *46th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*. 2005; 1897.

17. Engelbrecht AP. *Fundamentals of computational swarm intelligence*. John Wiley & Sons; 2006.
18. Vandenbergh F, Engelbrecht A. A study of particle swarm optimization particle trajectories. *Inf Sci*. 2006;176(8):937-971.
19. Hole AR, Yoo Hong II. The use of heuristic optimization algorithm to facilitate maximum simulated likelihood estimation of random parameter logit models. *Royal Stat Soc Appl Stat Ser C*. 2017;66:997-1013.
20. Sun J, Garibaldi JM, Hodgman C. Parameter estimation using metaheuristics in systems biology: a comprehensive review. *IEEE/ACM Trans Comput Biol Bioinf*. 2012;8(1):185-202.
21. Abdullah A, Deris S, Mohamad MS, Anwar S. An improved swarm optimization for parameter estimation and biological model selection. *PLoS One*. 2013;8(4):e61258.
22. Wang FK, Huang PR. Implenting particle swarm optimization algorithm to estimate the mixture of two Weibull parameters with censored data. *J Stat Comput Simul*. 2014;84(9): 1975-1989.
23. N, Nefti-Meziani S, Atyabi A. A comprehensive review of swarm optimization algorithms. *PLoS One*. 2015;10(5):e0122827.
24. Shikha A, Sanjay S. A review on application of particle swarm optimization in bioinformatics. *Curr Bioinform*. 2015;10(4):401-413.
25. Koç H, Dünder E, Gümüştekin S, Koç T, Cengiz MA. Particle swarm optimization-based variable selection in Poisson regression analysis via information complexity-type criteria. *Commun Stat-Theory Meth*. 2018;1:5298-5306.
26. Christian B, Raidl GR. *Hybrid Metaheuristics: Powerful Tools for Optimization*. Springer International Publishing; 2016.
27. Blum C, Puchinger J, Raidl GR, Roli A. Hybrid metaheuristics in combinatorial optimization: A survey. *Appl Soft Comput*. 2011;11:4135-4151.
28. Christian B, Grob R. *Swarm Intelligence in Optimization and Robotics*. Springer; 2015: 1291-1309.
29. Chen R-B, Chen P-Y, Hsu C-L, Wong WK. Hybrid algorithms for generating optimal designs for discriminating multiple nonlinear models under various error distributional assumptions. *PLoS One*. 2020;15(10):e0239864.
30. Shi Y, Wong WK, Goldin J, Brown MS, Kim HJ. Prediction of progression in idiopathic pulmonary fibrosis using quantum particle swarm optimization hybridized random forest. *Artif Intell Med*. 2019;100:101709.
31. Kim S, Li L. A novel global search algorithm for nonlinear mixed-effects models using particle swarm optimization. *J Pharmacokinet Pharmacodyn*. 2011;38:471-495.
32. Lindstrom M, Bates D. Nonlinear mixed effects models for repeated measures data. *Biometrics*. 1990;46:673-687.
33. Davidian M, Giltinan DM. Some simple methods for estimating intraindividual variability in nonlinear mixed effects models. *Biometrics*. 1993;49:59-73.
34. Beal SL, Sheiner LB. Estimating population kinetics. *CRC Crit Rev Bioeng*. 1982;8:195-222.
35. Vonesh EF, Carter RL. Mixed-effects nonlinear regression for unbalanced repeated measures. *Biometrics*. 1992;44:673-683.
36. Wolfinger RD. Laplace's approximation for nonlinear mixed models. *Biometrika*. 1993;80:791-795.
37. Proost JH, Schiere S, Eleveld DJ, Wierda JMKH. Simultaneous versus sequential pharmacokinetic-pharmacodynamic population analysis using an iterative two-stage Bayesian technique. *Biopharm Drug Dispos*. 2007;28:455-473.
38. Bulitta JB, Landersdorfer CB. Performance and robustness of the Monte Carlo importance sampling algorithm using parallelized S-ADAPT for basic and complex mechanistic models. *AAPS J*. 2011;13:212-226.
39. Bauer RJ, Guzy S. Monte Carlo parametric expectation maximization (MC-PEM) method for analyzing population pharmacokinetic/pharmacodynamic data. In: Dargenio DZ, ed. *Advanced Methods of Pharmacokinetic and Pharmacodynamic Systems Analysis Volume 3. The International Series in Engineering and Computer Science*. Springer; 2004:266-290.
40. Kim S, Hall SD, Li L. A novel Gibbs maximum a posteriori (GMAP) approach on Bayesian nonlinear mixed-effects population pharmacokinetics (PK) models. *J Biopharm Stat*. 2009;19(4):700-720.
41. Chan PLS, Jacqmin P, Lavielle M, McFadyen L, Weatherley B. The use of the SAEM algorithm in MONOLIX software for estimation of population pharmacokinetic-pharmacodynamic-viral dynamics parameters of maraviroc in asymptomatic HIV subjects. *J Pharmacokinet Pharmacodyn*. 2011;39:41-61.
42. Fidler M, Xiong Y, Schoemaker R, et al. nlmixr: Nonlinear mixed effects models in population pharmacokinetics and pharmacodynamics. R package version 2.0.1; 2021.
43. Liu X, Wang Y. Comparing the performance of FOCE and different expectation-maximization methods in handling complex population physiologically-based pharmacokinetic models. *J Pharmacokinet Pharmacodyn*. 2016;43:359-370.
44. Türken O, Tez M. An application of Nelder-Mead heuristic-based hybrid algorithms: Estimation of compartment model parameters. *Int J Artificial Intell*. 2016;14(1):112-129.
45. Nelder JA, Mead R. A simplex method for function minimization. *Comput J*. 1965;7:308-313.
46. Türken O, Tez M. Statistical inference for two-compartment model parameters with bootstrap method and genetic algorithm. In: Tez M, Rosen D, eds. *Trends and Perspectives in Linear Statistical Inference. Contributions to Statistics*. Springer; 2018.
47. Pan S, Zhou J, Zhou S, Huang Z, Meng J. Pharmacokinetic-pharmacodynamic modeling for Moutan Cortex/Moutan Cortex charcoal and the contributions of the chemical component using support vector regression with particle swarm optimization. *RSC Adv*. 2020;10:24454.
48. Toney MD. Common enzymological experiments allow free energy profile determination. *Biochemistry*. 2013;52(34):5952-5965.
49. Sowparnika GC, Thirumarimurugan M, Sivakumar VM, Vinoth N. Controlled infusion of intravenous cardiac drugs using global optimization. *Indian J Pharmacol*. 2019;51(1):61-71.
50. Arijit B, Sambarta D, Swagatam D, Ajith A. *Synergy of PSO and Bacterial Foraging Optimization – A Comparative Study on Numerical Benchmarks*. Springer, Berlin Heidelberg; 2007: 255-263.
51. Chen Y-P, Li Y, Wang G, et al. A novel bacterial foraging optimization algorithm for feature selection. *Expert Syst Appl*. 2017;83:1-17.
52. Ye L, Liu J, Wang Y, et al. Development of a three-compartment toxicokinetic model for T-2 toxin in shrimp by blindfold particle swarm optimization algorithm. *Ecotoxicol Environ Saf*. 2021;208:111698.
53. Godfrey KR, Fitch WR. The deterministic identifiability of nonlinear pharmacokinetic models. *J Pharmacokin Biopharm*. 1984;12:177-191.

54. Kim S, Li L. A switching Markov chain Monte Carlo method for statistical identifiability of nonlinear pharmacokinetics models. *Stat Sin*. 2012;22(3):1199-1215.
55. S, Li L. Statistical identifiability and convergence evaluation for nonlinear pharmacokinetic models with particle swarm optimization. *Comput Methods Programs Biomed*. 2014;113(2):413-432.
56. Godfrey KR, DiStefano JJ. *Identifiability of Model Parameters in Identifiability of Parametric Models*. Pergamon Press; 1987.
57. Vajda S, Godfrey KR, Rabitz H. Similarity transformation approach to identifiability analysis of nonlinear compartmental-models. *Math Biosci*. 1989;93:217-248.
58. Lecourtier Y, Lamnabhi-Lagarrigue F, Walter E. *Volterra and generating power series approaches to identifiability testing, in identifiability of parametric models*. Pergamon Press; 1987.
59. Saccomani MP, Audoly S, D'Angiò L. Parameter identifiability of nonlinear systems: the role of initial conditions. *Automatica*. 2003;39:619-632.
60. Hengl S, Kreutz C, Timmer J, Maiwald T. Data-based identifiability analysis of non-linear dynamical models. *Bioinformatics*. 2007;23:2612-2618.
61. Shivva V, Korell J, Tucker IG, Duffull SB. An approach for identifiability of population pharmacokinetic pharmacodynamic models. *CPT: Pharm Syst Pharmacol*. 2013;2(6):e49.
62. Aoki Y, Nordgren R, Hooker AC. Preconditioning of nonlinear mixed effects models for stabilisation of variance-covariance matrix computations. *AAPS J*. 2016;18(2):505-518.
63. Bjugård Nyberg H, Hooker AC, Bauer RJ, Aoki Y. Saddle-reset for robust parameter estimation and identifiability analysis of nonlinear mixed effects models. *AAPS J*. 2020;22(4):90.
64. Hidalgo ME, Ayesa E. A numerical identifiability test for state-space models – application to optimal experimental design. *Water Sci Technol*. 2001;43:339-346.
65. Wynn HP, Parkin N. Sensitivity analysis and identifiability for differential equation models. *Proceedings of the 40th IEEE Conference on Decision and Control (Cat. No.01CH37228)*. 2001; 3116-3121.
66. Yue H, Brown M, He F, Jia J, Kell DB. Sensitivity analysis and robust experimental design of a signal transduction pathway system. *Int J Chem Kinet*. 2008;40:730-741.
67. Lindley DV. *Bayesian Statistics: A Review*. SIAM; 1987.
68. Eberly LD, Carlin BP. Identifiability and convergence issues for Markov chain Monte Carlo fitting of spatial models. *Stat Med*. 2000;19:2279-2294.
69. Li J, Kim S, Shields AF, et al. Integrating dynamic positron emission tomography and conventional pharmacokinetic studies to delineate plasma and tumor pharmacokinetics of FAU, a prodrug bioactivated by thymidylate synthase. *J Clin Pharmacol*. 2016;56(11):1433-1447.
70. Chernoff H. Locally optimal designs for estimating parameters. *Ann Math Stat*. 1953;24(4):586–602.
71. Atkinson AC, Donev AN, Tobias RD. *Optimum Experimental Designs, with SAS*. Oxford Statistical Science, Oxford University Press; 2009.
72. Fedorov VV, Leonov SL. *Optimal Design for Nonlinear Response Models*. Chapman and Hall/CRC Press; 2014.
73. Radhakrishna RC. Information and accuracy attainable in the estimation of statistical parameters. *Bullet Calcutta Math Soc*. 1945;37(3):81-91.
74. Harald C. *Mathematical methods of statistics*. Princeton University Press; 1946.
75. Mentré F, Mallet A, Baccar D. Optimal design in random effect regression models. *Biometrika*. 1997;84:429-442.
76. Mielke T, Schwabe R. Some considerations on the Fisher information in nonlinear mixed effects models. Proceedings of the 9th International Workshop in Model-Oriented Design and Analysis 2010:129-136.
77. Nyberg J, Bazzoli C, Ogungbenro K, et al. Methods and software tools for design evaluation in population pharmacokinetics-pharmacodynamics studies. *Br J Clin Pharmacol*. 2015;79(1): 6-17.
78. Bazzoli C, Retout S, Mentré F. Fisher information matrix for nonlinear mixed effects multiple response models: evaluation of the appropriateness of the first order linearization using a pharmacokinetic/pharmacodynamic model. *Stat Med*. 2009;28(14):1940-1956.
79. Hooker A, Vicini P. Simultaneous population optimal design for pharmacokinetic-pharmacodynamic experiments. *AAPS J*. 2005;7:759-785.
80. Nguyen TT, Bazzoli C, Mentré F. Design evaluation and optimisation in crossover pharmacokinetic studies analysed by nonlinear mixed effects models. *Stat Med*. 2012;31(11–12):1043-1058.
81. Riviere M-K, Ueckert S, Mentré F. An MCMC method for the evaluation of the Fisher information matrix for non-linear mixed effect models. *Biostatistics*. 2016;17:737-750.
82. Ueckert S, Mentré F. A new method for evaluation of the Fisher information matrix for discrete mixed effect models using Monte Carlo sampling and adaptive Gaussian quadrature. *Comput Stat Data Anal*. 2017;111:203-219.
83. Plumlee M. Fast prediction of deterministic functions using sparse grid experimental designs. *J Am Stat Assoc*. 2014;109:1581-1591.
84. Retout S, Comets E, Samson A, Mentré F. Design in nonlinear mixed effects models: optimization using the Fedorov-Wynn algorithm and power of the Wald test for binary covariates. *Stat Med*. 2007;26(28):5162-5179.
85. Foracchia M, Hooker A, Vicini P, Ruggeri A. poped, a software for optimal experiment design in population kinetics. *Comput Meth Progr Biomed*. 2004;74(1):29-46.
86. Nyberg J, Ueckert S, Strömberg EA, Hennig S, Karlsson MO, Hooker AC. PopED: an extended, parallelized, nonlinear mixed effects models optimal design tool. *Comput Meth Progr Biomed*. 2012;108(2):789-805.
87. Shi YU. Using particle swarm optimization to find optimal longitudinal designs for mixed effects models with sparse grid and predict progression of idiopathic pulmonary fibrosis using single time point high resolution computed tomography scans with random forest. *PhD Dissertation, UCLA Department of Biostatistics*; 2019. https://escholarship.org/uc/item/0k99z04n.
88. Shi Y, Eberhart RC. Empirical study of particle swarm optimization. *Proceedings of the 1999 Congress on Evolutionary computation-CEC99 (Cat. No. 99TH8406)* 1999; 3: 1945-1950.
89. Tong XT, Choi KP, Lai TL, Wong WK. Stability bounds and almost sure convergence of improved particle swarm optimization methods. *Res Math Sci*. 2021;8:30.
90. Lai TL, Choi KP, Tong TX, Wong WK. A statistical approach to adaptive parameter tuning in nature-inspired optimization and optimal sequential design of dose-finding trials. *Statistica Sinica*. 2021;31:1-21.

91. Tsai C-W, Chiang M-C, Ksentini A, Chen M. Metaheuristic algorithms for healthcare: Open issues and challenges. *Comput Electr Eng*. 2016;53:421-434.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Kim S, Hooker AC, Shi Y, Kim GHJ, Wong WK. Metaheuristics for pharmacometrics. *CPT Pharmacometrics Syst Pharmacol*. 2021;10:1297–1309. doi: 10.1002/psp4.12714