Efficient and flexible simulation-based sample size determination for clinical trials with multiple design parameters

Context: Simulation offers a simple and flexible way to estimate the power of a clinical trial when analytic formulae are not available. The computational burden of using simulation has restricted its application to only the simplest of sample size determination problems, minimising the overall sample size subject to power being above a target level. Many sample size determination problems are more complicated. There can be several design parameters over which to optimise and several conflicting criteria to simultaneously minimise. As the complexity of the sample size determination problem increases, so too does the computational burden incurred by using simulation to estimate power. Efficient global optimisation algorithms for the design and analysis of computer experiments could help manage this burden, enabling simulation-based sample size determination for complex problems.

Objective: To develop and illustrate a framework for using efficient global optimisation techniques to solve complex simulation-based sample size determination problems.

Methods: We describe a general framework for solving multi-parameter multi-criteria sample size determination problems based around an established global optimisation algorithm. The method involves using a non-parametric regression model as an approximation of the true underlying power function. The model predictions of power at yet-to-be evaluated sample sizes are then used at each iteration to carefully select the point to be evaluated next. The method is flexible, can be used for almost any problem for which power can be estimated using simulation, and can be implemented using existing statistical software packages.

Results: We use the framework to revisit the design of a large trial of secondary care treatments for chronic fatigue syndrome with treatment-related clustering. Sample size determination involves simultaneously minimising the numbers of participants, therapists and doctors. We show how the method can be used to find a set of designs with nominal power and offering different balances between these three minimisation criteria.

Conclusions: Efficient global optimisation based on surrogate models can be an effective method for complex sample size determination problems when power must be estimated using simulation. By improving the efficiency of these calculations, increasingly complex problems can be addressed in a timely manner with a common method.