Model simulation and decision analysis with the SimR package in R

Joseph Kahn*, Peter Danenberg, Sourav Das, Derek Ayers, Richard Nixon, Joyee Ghosh
Novartis Pharmaceuticals, East Hanover, NJ, USA.

Objectives: To better enable modeling and simulation to inform clinical development decisions, we developed an automated simulation package in R incorporating good practices from decision analysis: the package enables quantitative comparison of decisions; it determines influential uncertainties; and it maintains transparency regarding source information and probabilistic dependencies.

Methods: Performing model simulation and decision analysis with SimR requires:

1. Assigning model input parameters and meta-data (e.g., units, value ranges, prior probability distributions and source/pedigree), along with the selection of decisions/strategies. These are stored in a spreadsheet or in .csv files for easy editing and review;
2. Writing an R function to carry out a single model run, (the only code required to be written and validated by the user);
3. Calling the SimR package in R that plans and runs simulations using appropriate random seeds, archives results, and produces analysis tables and plots.

Results: Multiple decision analysis projects are being carried out at Novartis pharmaceuticals in R with the SimR package. For increased speed, the package can run parallel simulations in a grid environment. "Tornado" plots of one-way sensitivity help determine which input parameters are most salient. Two-way sensitivity plots display interaction effects from input parameters to model outputs. Bar charts and cumulative distribution plots compare performance and risks among strategies.

Conclusions: Modeling and simulation using the SimR package is grounded in decision analysis best practices. Its benefits include:

1. Faster model development and validation with use of functions for bookkeeping and production of analysis tables and graphics;
2. Facilitation of rapid communication and model archiving with standardized inputs and results formats that free users to focus on insightful comparisons;
3. Fast run times and greater precision with multiple simulations automatically run in parallel on a grid computing network.

References:

[1] Ross Ihaka and Robert Gentleman (1996). "R: A Language for Data Analysis and Graphics," *Journal of Computational and Graphical Statistics*, **5** (3): 299—314.