Abstracts of the Ninth Workshop on Simulation (2018) 1-2

Performance measures in dose-finding experiments

Nancy Flournoy [1], José A. Moler [2], Fernando Plo [3]

Consider binary responses indicating toxicity for which the probability of toxicity is assumed to increase with dose. In the first phase of pharmaceutical development in oncology, the main statistical goal is to estimate a given target percentile from such binary responses. In this work, we present tools for studying the performance of adaptive designs in this context, taking measures of inferential precision and toxic exposure into account simultaneously. Finally, these tools are used in a simulation study that evaluates the performance of selected phase I procedures.

1 Introduction

In this work, we study combined measures of ethical treatment and quality estimation to clarify their tradeoff and to quantify the global performance of a design. Then, selected competing designs are compared with each other with respect to these measures. Of course, these measures depend on the method of quantile estimation and/or dose-selection used.

By simulation, summary ethical and inferential measures are obtained for each n, as n increases over a preset range. The Pareto frontier is graphically obtained from a plot of a selected ethical and inferential measure. Such graphs are shown to be useful for studying the performance of various allocation rules with respect to both criteria simultaneously.

For an ethical criterion, many papers in the literature use some measure of how close dose allocations are to the target dose. These measures are surrogates for patients' toxic exposure. In a simulation study, toxic exposure can be measured directly. Therefore, we measure the ethics of a design in terms of the (simulated) observed toxicity rates. Measures of how close doses are to the target dose we call *allocation criteria*. Procedures can perform identically with respect to studies overall toxicity rates while having quite distinct dose-allocation patterns.

We study several measures of the quality of estimators. First, the root mean square error is measured on the dose scale, that is, using the distance of each quantile estimator from the target dose. It is also measured on the toxicity scale, that is, using the distance between the simulated toxicity rate and actual toxicity rate at the target dose. Finally, in order to provide a performance measure in terms of patients (to match the ethical criterion), the classical optimal design concept of efficiency is included.

[1]University of Missouri Curators Distinguished Professor Emerita, 600S State St., 408 Bellingham, WA 98225,E-mail: flournoyn@missouri.edu

[2]Department of Statistics, Public University of Navarre, Pamplona, 31015, Spain, E-mail: jmoler@unavarra.es

[3]Department of Statistics, University of Zaragoza, Pedro Cerbuna, 12, Zaragoza, 50009, Spain, E-mail: fplo@unizar.es

References

[1] Oron A. and Hoff P. *Small-sample behavior of novel phase I cancer trial designs* //Clinical Trials, 2013, v. 10, p. 63–80.

[2] Stylianou M. and Flournoy N. *Dose finding using the biased coin up-and-down design and Isotonic Regression.* //Biometrics, 2002, v. 58, p. 171–177.

[3] Sverdlov O., Wong W.K. and Ryeznik Y. *Adaptive clinical trial designs for phase I cancer studies*// Statistical surveys, 2014, v. 8, p. 2–44