CRM Method Description

John D. Cook

July 15, 2005

1 Introduction

This document gives the statistical background for the CRM method as implemented in this software application. Other implementations offer more generality but are harder to use; this implementation optimized for ease of use rather than statistical generality.

2 The goal of CRM

The goal of the CRM method is to determine which of a finite list of doses has probability of toxicity closest to a specified target. For example, one may wish to find the dose that is toxic to one out of every four patients. See O'Quigley *et al* 1990.

The CRM method rests on the assumption that the response due to the agent being tested increases with dose. For this reason, one would like to give as much of the agent as possible. However, it is also assumed that toxicity increases with dose and that toxicity must limit dose. Therefore one aims to find the "maximum tolerated dose" or MTD, the dose with probability of toxicity closest to some target.

Note that the CRM method only uses toxicity data; effectiveness plays no role. If the lowest dose is effective but not sufficiently toxic, CRM will seek out a higher dose. If your goal is to find an *effective* dose while minimizing toxicity, you may be interested in the EffTox method. See Thall and Cook 2004.

3 Governing accrual

Patients in a CRM trial are treated in cohorts, groups of patients who receive the same dose. The most common cohort size is three, though one could set the cohort size to 1, effectively eliminating cohorts. The outcomes of all patients in a cohort must be observed before calculating the recommended dose for the next cohort.

A stopping rule is included to protect against the possibility that even the lowest dose is too toxic. If the posterior probability that the lowest dose is more toxic that target is above a specified threshold value, say 0.9, then the trial will stop. To eliminate the stopping rule, the threshold value may be set to 1. However, this would be unethical in an actual trial because it would imply that the trial never stops, even if all patients experience toxicity.

If patients are treated in cohorts of size c, then up to c patients may be treated at a toxic dose level before de-escalating or stopping the trial. Therefore c must not be too large. This implementation requires $c \leq 4$.

4 Probability model

The probability of toxicity at dose i is modeled as $p_i^{\exp(\alpha)}$ where p_i is a constant and α is distributed *a priori* as a normal random variable with mean 0 and variance 2.

In an earlier version of CRM, we allowed the prior variance to be specified by the user. There are four reasons why we now fix the value of this parameter.

1. Experience showed that changing this parameter did not significantly effect operating characteristics.
2. Nearly all users set the prior variance to 2.
3. The software was made more robust and more efficient by being able to assume a fixed prior variance.
4. The current implementation of CRM emphasizes simplicity rather than generality.

The CRM method requires that the investigator specify his or her prior mean probability of toxicity at each of the doses under consideration. Denote these values $s_1, s_2, \ldots, s_n$. The set of s values is called the "skeleton" of the CRM. Because the CRM model assumes that toxicity increases with dose, the s_i values must increase as i increases. This implementation further constrains the skeleton values so that $0.01 \geq s_1$ and $s_n \leq 0.99$.

The values p_i in the probability model are selected so that $E[p_i^{\exp(\alpha)}] = s_i$. In other words, the method solves for the values p_i so that the prior mean probability of toxicity at each dose is the elicited value.

The dose given to the first cohort is chosen arbitrarily and not based on the probability model. Often this value is chosen to be the lowest dose due to safety concerns. One may believe *a priori* that a higher dose has probability of toxicity closer to the target, but choose to start at a lower dose. However, one need not start at the lowest dose. After all, one's choice of "lowest" dose is arbitrary since the list of doses to consider is arbitrary.

After the first cohort, each successive cohort is given the dose whose posterior probability of toxicity given the data collected thus far is closest to the target, subject to one additional requirement: one cannot skip over an untried dose. If the method would otherwise skip over an untried dose, the lowest untried dose is given instead. If the method does not start at the lowest dose, all doses below the starting dose are considered "tried" for purposes of the no-skip rule. For example, if a trial of 5 dose levels starts with dose 3, the method could give the second cohort dose 1, 2, 3, or 4. But if the method determined that dose 5 had posterior probability of toxicity closest to the target, the software would assign dose 4 instead.

Note that the final decision of the trial is *not* the dose that was given to the last cohort, though it could be. The final outcome of the trial is the dose that *would* be given the next cohort if there were one. Otherwise, the outcomes of the last cohort would effectively be thrown away.

5 Simulations

In order to understand the operating characteristics of a CRM trial, one may simulate how the trial behaves under various scenarios. Each scenario specifies a set of "true" probabilities of toxicity at each dose. One then asks what would happen, on average, if reality corresponded to these hypothetical values. This is done by repeating the trial many times, say 100 times, and observing how often each dose is given or selected as the final dose on average.

One's choice of scenarios is arbitrary. However, one typically considers at least one bad scenario and one good scenario, as well as a few in between. A "bad" scenario is one in which the lowest dose is quite toxic. Here one wants to verify that the method stops sufficiently often. In a "good" scenario, one of the doses has probability of toxicity close to the target, and one wants to verify that the method seldom stops and often finds the best dose.

It is also informative to run a scenario using the skeleton probabilities as the scenario probabilities. These skeleton values represent the best guess at what the trial will reveal. If the method doesn't perform well when it does exactly what one expects, the design needs to be revised.

6 References

O'Quigley, J., Pepe, M., Fisher, L. (1990). Continual reassessment method: a practical design for phase I clinical trials in cancer. *Biometrics* **46**, 33-48.

Thall, Peter F., Cook, John D. (2004). Dose-Finding Based on Efficacy-Toxicity Trade-Offs. *Biometrics* **60**, 684-693.