Cancer Chemother Pharmacol (1998) 41: 429–436

ORIGINAL ARTICLE

S. Piantadosi · J.D. Fisher · S. Grossman

Practical implementation of a modified continual reassessment method for dose-finding trials

Received: 6 May 1997 / Accepted: 25 July 1997

Abstract *Purpose*: We describe a practical, reliable, efficient dose-finding design for cytotoxic drugs applied in a multi-institutional setting. *Methods*: The continual reassessment method (CRM) was modified for use in phase I trials conducted through the New Approaches to Brain Tumor Therapy (NABTT) Consortium. Our implementation of the CRM uses (1) a simple dose-toxicity model to guide data interpolation, (2) groups of three patients to minimize calculations and stabilize estimates, (3) investigators' clinical knowledge or opinion in the form of data to make the process easier to understand, and (4) a flexible computer program and interface to facilitate calculations. *Results*: The modified CRM was used in two dose-finding trials of 9-aminocamptothecin in patients with newly diagnosed and recurrent glioblastoma who were taking anticonvulsant medication. The CRM located the maximum tolerated dose (MTD) efficiently in both trials. Compared to conventional designs, the CRM required slightly more than half the number of patients expected, did not greatly overshoot the MTD (i.e. no patients were treated at dangerously high doses), and did not underestimate the MTD. *Conclusions*: Our experience demonstrates the feasibility of implementing this design in multi-institutional trials and the possibility of performing dose-finding studies that require fewer patients than conventional methods.

Key words Continual reassessment method · Phase I trials · Standard method (SM) · Glioblastoma · 9AC

This work was supported by NIH/NCI Grants CA62474 and CA62475

S. Piantadosi · J.D. Fisher · S. Grossman
Johns Hopkins Oncology Center, Baltimore, Maryland, USA

S. Piantadosi (✉)
NABTT CNS Consortium Central Operations Office,
600 North Wolfe Street, Baltimore, Maryland 21287, USA

Introduction

Finding optimal doses of new cytotoxic drugs, or extending the use of existing agents into new patient populations, requires a practical, reliable, efficient dose-finding design. The best therapeutic index for most cytotoxic drugs is obtained using a relatively high dose that yields manageable side effects, the so-called "maximum tolerated dose" (MTD). Some patients would be able to tolerate higher doses, albeit with increasing incidence and severity of side effects or toxicities. Dose-finding study designs do not select a "maximum", but an "optimum", i.e. a dose that satisfies constraints imposed by the practitioner. We refer to the optimum dose as the "target".

Estimating the target dose efficiently can be a challenge, even in a homogeneous cohort of cancer patients with good function in major organ systems. Ethical considerations require building clinical evidence beginning with low doses and progressing to high doses, which is usually not the optimal experimental design. Doses are modified according to the frequency of toxicity, which is imprecisely measured in severity and time. Clinicians prefer operationally simple designs with definitive decision rules and small sample sizes for determining the optimal dose. It is usually not necessary to select doses with a high degree of precision, but a reliable answer is important. Additional dose finding in later development of the drug is expensive and time consuming.

Dose-finding (phase I) clinical trials attempt to meet these challenges using designs that have evolved from bioassay experiments and are familiar to oncologists. An excellent new dose-finding design, the continual reassessment method (CRM), has been suggested in recent years [1]. It continues to be studied and improved in the statistical literature, even though it has not been widely applied by oncologists. After a simulation study, Korn et al. [2] recommended a standard dose-finding method over the CRM. However, Goodman et al. [3] successfully modified the CRM to address all the shortcomings noted by Korn et al. Despite its potential advantages,

the CRM has been slow to diffuse into practice because of some practical difficulties (discussed below).

In this paper we present a workable implementation of the CRM, and relate our clinical experience with it. The principal improvements that we suggest are employing (small) groups of patients at each dose, representing investigator knowledge in the form of data throughout the dose-finding process, frequently revising the quantitative aspects of dose finding as new information becomes available, and employing a friendly interface for computations. We begin by reviewing some widely used dose-finding designs and the CRM. We then discuss a specific implementation of the CRM that has proved to be feasible in the New Approaches to Brain Tumor Therapy (NABTT) consortium, an NIH funded group of collaborating institutions performing early developmental trials in central nervous system malignancies. Finally, we present specific examples where the CRM has been used and improved the efficiency and quality of clinical inferences.

Background

Dose-finding designs commonly used

The most widely used dose-finding designs have been variations of the "up-and-down method" [4, 5]. That design begins with a starting dose and specification of an ordered set of higher doses to be tested, one of which will be selected as the MTD. The dose employed in the next patient or group of patients is the next higher or lower dose, depending on outcomes at the current dose and simple decision rules. An example of a set of decision rules might be: (1) three patients are treated at each dose, (2) if no more than one serious toxicity is seen, the dose increases, (3) if two or more toxicities are seen the dose decreases. The design terminates when the frequency of toxicity (or other response) meets or exceeds a specified level.

Many variations of this type of basic design have been used in dose escalation studies (reviewed, with one example, in reference 6). The stochastic approximation method [7] is a little used, but interesting, method. Experimental design issues for dose response studies are discussed by several authors including Wong and Lachenbruch [8], Ruberg [9, 10], and Ratain et al. [11]. Most dose-finding designs applied to human trials are derived from the up-and-down method and have features in common. These are:

1. The starting dose is based on preclinical data, the action of similar drugs, patient factors, and investigator judgment.
2. The investigator prespecifies a small set of candidate doses to be employed during the trial, one of which will be selected as the target, a dose-escalation plan, and semiquantitative clinical decision rules for study termination.
3. The clinical investigator formally assesses outcomes and makes decisions about dose increases or decreases.
4. Tiny cohorts of patients are treated at each dose (e.g. one to three), on which to base decisions.
5. Results from distant doses do not quantitatively or statistically affect the assessment of results from the current dose.
6. The target tends to be underestimated because of patient variability, the sequential nature of the basic design, and the discreteness of the dose set.

As originally designed, the up-and-down method employed a dose-response model after all data were obtained to facilitate estimating the target. However, the designs used most frequently have abandoned this final modeling step, and define the target in terms of the last dose employed. Thus, these designs select a target dose operationally.

CRM design

The CRM is fundamentally different from widely used implementations of the up-and-down design and its variants. There are two essential differences that lead to improvements in points 2, 5, and 6 discussed above. First, the CRM attempts to estimate the target from a continuum of doses, whereas operational designs merely select a dose from a discrete set. If the true target is not among the choices set out in advance by operational designs, they can only approximate it.

The second essential difference is that the CRM uses a mathematical model for dose-response to summarize the accumulating data and to guide selection of the next dose. Data become available when the first patient or cohort is treated at the starting dose but also may be available beforehand from other studies. The dose-response model is fitted to the data and used to estimate the target dose. The updated estimate can be higher or lower than a previously used dose and is not constrained to be one of a set of doses specified at the start of the experiment. The new dose is then utilized in the next patient or cohort. The steps of treating new patients, data gathering, model fitting to all the data, and dose updating repeat until the process converges, i.e. until the dose no longer changes.

In addition to these points, the CRM relies on a direct quantitative connection between a response, such as toxicity, and a dose. However, the CRM permits using drug concentration, peak concentration, area under the curve (AUC), or any other meaningful pharmacokinetic parameter in place of dose (or in addition to dose) as the determinant of response [12]. In fact, for simple compartmental models of drug distribution, AUC and dose are proportional (see, for example, reference 13).

The fact that the CRM uses a dose-response model (but not the specific model chosen) creates the strengths and weaknesses of the method. Therefore, it is worth briefly discussing the modeling process, which requires

art as much as science. When we postulate a model, we specify only its mathematical structure, leaving one or more constants or parameters to be determined by actual data. To be most useful, the model must be realistic, parsimonious, and flexible. *Realism* means that the model can behave qualitatively only in the correct way. *Parsimony* means that the model does not contain unnecessary complications. *Flexibility* means that the model can represent important nuances in real data.

Suppose the relationship between drug dose and the probability of response is a sigmoid curve. This may be reasonable for toxicities related to cytotoxic drugs, although it is not universally true. A model that behaves in the correct way could increase the efficiency of observation by representing this knowledge. If we have only a few observations, the model may contribute substantially to the available information. If we have a large number of observations, the model may be the best tool for summarizing data. Using the model is motivated by three unverifiable assumptions: (1) the true dose-response relationship has a certain form (biological), (2) we can specify a mathematical model that would mimic observation if large amounts of data were collected (empirical), and (3) a model, even an imperfect one, can capture and represent biological knowledge (epistemological).

To illustrate a specific model, assume that the true dose-toxicity function is a logistic curve such that the chance of a grade 3 or 4 toxicity is

$$\Pr[\text{toxicity}] = \frac{1}{1 + e^{-\beta(d - d_{50})}}, \qquad (1)$$

where d is the dose employed, β is a steepness parameter, and d_{50} is the dose associated with half maximal response (i.e. when $d = d_{50}$, the probability of toxicity is $\frac{1}{2}$). The pedigree for this particular model, which can also be written in other ways, is based mainly on bioassay, laws of mass action, and mathematical convenience. The logistic function looks like Fig. 1 where the horizontal location and steepness of the curve can change according to β and d_{50}. This model satisfies the requirements outlined above.

If we estimate the parameters, denoted by $\hat{\beta}$ and $\hat{d}_{50}$, by fitting this curve to data, we can calculate the dose associated with any probability of toxicity by solving Eq. 1 for d. For example, the dose that yields a 30% chance of toxicity is

$$d_{30} = \hat{d}_{50} + \log\left\{\frac{0.3}{1 - 0.3}\right\} \Big/ \hat{\beta} \ . \qquad (2)$$

When incorporated into dose-finding, a model-based approach yields important advantages over other methods:

1. The model captures biological knowledge that is relevant to the problem, making very efficient use of the data.
2. The method yields unbiased estimates of the target dose.

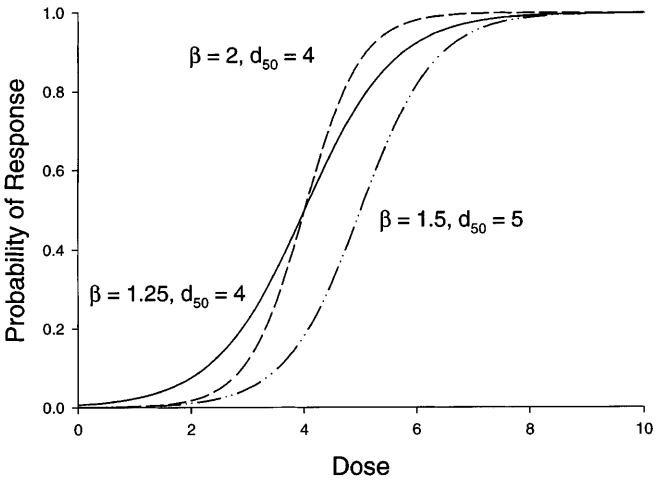

Fig. 1 Logistic dose-toxicity function. The dose scale is arbitrary

3. A fixed set of doses does not have to be specified in advance of the trial.
4. Patient data from outside the dose escalation can be utilized formally.
5. Changing clinical impressions can be incorporated in the dose escalation.
6. The recommended dose can be updated after the last group of patients has been treated.
7. The model can be used to adjust the dose escalation for clinical or pharmacological parameters in addition to, or other than, dose (e.g. AUC).

Most of these advantages will be illustrated below.

Methods

Based on the previous discussion, it is evident how the CRM sustains a search for the target as data become available. It is less obvious how the process gets started. Initiating the CRM requires information to use in the dose-response model. Because little or no data are available at the outset, this information must come from the clinician. The initial information can come in any one of several forms: (1) observations of patients, (2) explicit quantitative specification of a model, (3) assumed probability distributions for the model parameters, (4) investigator judgment formalized as "data", or (5) a mixture of the above.

Investigators accustomed to informal methods of choosing a starting dose may worry that the information required to initiate the CRM is created from nothing. However, information is not created but merely summarized in a different usable form. For example, consider the typical phase I design where investigators specify the starting dose and a set of doses to be tried in the trial. The same information used to construct this scheme can be used to initiate the CRM. Furthermore, in the usual dose escalation, one of the prespecified doses will be chosen as the MTD. We cannot expect the clinician to name the MTD reliably before the experiment is conducted. The CRM is free of this problem.

Some information about both high and low doses is needed to facilitate the model fit. As data from low doses become available, the model can rely on them. However, the need for investigator input into the model for high doses may continue throughout the CRM steps. This appears to be quite feasible, although it can be difficult in the presence of large patient variability. The same circumstance presents problems for traditional dose-finding methods.

Our suggestions for how to initiate and continue the CRM are discussed in the following points.

1. If no ancillary information is to be used formally at the start of the dose escalation, determine the starting dose in the usual ways and proceed to step 6.
2. *Use preclinical and other information about low doses*. Information about responses (toxicities) may be available from a previous study. For example, if patients receive the drug but it produces no toxicity, a dose-finding trial may be initiated and the preexisting data will be helpful in determining the starting dose. The drug or its analogs may have been tested in a different population, which could yield quantifiable information.
3. *Formalize and quantify clinical intuition about drug behavior at high and low doses*. It is most helpful for the clinician to provide two pieces of information for the model in the form of data: a point estimate of response probability at a specific dose and a degree of certainty in that estimate. For example, a point estimate can be obtained by asking what dose would most likely yield a 10% chance of response (d_{10}). The degree of certainty about d_{10} can be quantified in the form of a numerator and denominator, e.g. 1 response out of 10 (not very certain) or 10 responses out of 100 (very sure). The same information is needed for a high dose, e.g. d_{90}. Convenient points other than 90% and 10% at each end of the dose-response range could be used. To simplify model fitting, investigators should choose points where the point estimate is neither zero nor one, i.e. the numerator is neither zero nor equal to the denominator. This method uses data to embody both clinical evidence (or judgment) and strength of evidence. The benefits of this approach are that

   (a) the need for, and nature of, assumptions is overt,
   (b) assumptions take the form of data which investigators understand and the model can use, and
   (c) data are used to initiate the model fitting in lieu of assumptions about the model parameters.

4. *Fit the dose-response model to the data*. The required data are triplets of numbers representing dose, number of patients treated, and number of responses. A convenient and widely employed method of model fitting (or equivalently, parameter estimation) uses the likelihood function, which expresses the probability of observing the data, given the model, as a function of the unknown parameters. The likelihood is based on the binomial probability distribution, reflecting the fact that r_i out of n_i responses are observed at dose$_i$. The values of d_{50} and β that maximize the likelihood are the "best" representations of the data. See Appendix I for computational details. At least one data point is required for each model parameter to be estimated. The computations for model fitting could be accomplished in a variety of ways. We have written a computer program (Appendix II) to perform the calculations in a rapid, flexible fashion. This high level language has capabilities, syntax, and a user interface well suited to such tasks.
5. *Invert the fitted model to estimate the target dose*. This is accomplished using Eq. 2.
6. *Gather data from patients using the current target dose*. We recommend that at least three patients be treated routinely at each dose. This number should be increased or decreased if investigators believe that the clinical circumstances demand it.
7. *Revise estimates of the dose extreme* d_{90}. This step is a subjective but critical reassessment of investigator judgment. As evidence accumulates at lower doses, investigators should reconsider what dose is likely to yield a 90% chance of response. Once actual data are accumulated at low doses, the investigator's guess about d_{10} is no longer needed.
8. Repeat steps 3–6 until the target dose changes by less than 10% or meets some other appropriate tolerance criterion.
9. Use the target dose for future developmental trials.

Results

To illustrate the flexibility and power of the CRM, we discuss applications developing cytotoxic drugs for patients with malignant gliomas. The drug being investigated was 9-aminocamptothecin (9-AC), a promising new antineoplastic agent [14]. 9-AC is an analog of the topoisomerase I inhibitor, camptothecin, which is an alkaloid extracted from the Chinese plant *Camptotheca acuminata*. Early studies of this drug demonstrated that it inhibits RNA and DNA synthesis in various human and animal cell lines in vitro and in vivo.

Dose escalation trials were initiated in four groups of patients: those with newly diagnosed cancer and those with recurrent disease, with and without anticonvulsants. All patients provided informed consent for participation and were treated as part of research studies approved by local Institutional Review Boards. Anticonvulsants are strong inducers of the hepatic P450 system that enhances metabolism of many drugs. The two dose escalations in patients without anticonvulsants have not concluded. All escalations employed a logistic dose-response function.

Patients with newly diagnosed malignant gliomas

Patients eligible for this trial were adults with newly diagnosed glioblastoma multiforme (GBM) who were taking anticonvulsants. Originally, patients were entered on a safety/efficacy (phase II) trial of 9-AC at a fixed dose of 850 $\mu g/m^2$ per day for 3 days as a continuous infusion. When it became clear that individuals on this trial, and those in a companion study of patients with recurrent anaplastic astrocytoma or GBM, were not experiencing clinically significant toxicity at the MTD, the safety/efficacy trial was terminated and a dose escalation was started. The clinical setting and details of this trial are reported elsewhere [15].

Initiation

The initial data showed 0 out of 22 patients (newly diagnosed patients plus those with recurrent disease) experienced dose-limiting toxicity (DLT) at 850 $\mu g/m^2$. This was convincing evidence that additional dose finding was needed, and we began by using the logistic model to guide our selection of the first CRM dose. Because the postulated dose-response model is never exactly zero, we used the one-sided upper 95% exact binomial confidence limit on 0/22 responses and assumed that the dose-response curve passed through the point (850, 0.13). This assumption was conservative in the sense that the response probability (0.13) was taken to be higher than the point estimate from the data (0.0).

To stabilize the right-hand portion of the model, we assumed that $d_{50} \approx 1600$. This was crudely in accord

Table 1 Values of β and MTD_{30} calculated for different d_{50} values

$\hat{d}_{50}$	$\hat{\beta}$	d_{30}
1350	0.0038	1127
1400	0.0035	1158
1500	0.0029	1208
1600	0.0025	1261
1700	0.0022	1315
1800	0.0020	1376
1900	0.0018	1429

with the existing data and clinical intuition but could not be verified. These assumptions allowed us to calculate $\hat{\beta}$:

$$0.13 = \frac{1}{1 + e^{-\hat{\beta}(850-1600)}}$$

$$\hat{\beta} = \frac{-\log\left(\frac{1}{0.13} - 1\right)}{850 - 1600} = 0.0025$$

The value found for $\hat{\beta}$ was not very sensitive to the assumed value for $\hat{d}_{50}$ (Table 1). Using these parameters, the starting dose satisfied

$$0.3 = \frac{1}{1 + e^{-0.0025(\hat{d}_{30}-1600)}}$$

or $d_{30} = 1261$. Thus, the first dose used in the CRM was $1260\,\mu g/m^2$.

Subsequent steps

Four patients were treated at a dose of $1260\,\mu g/m^2$. The results and subsequent steps are shown in Table 2 where n denotes the number of patients treated and r denotes the number of responses (toxicities). Several refinements were made during the iterations. These illustrate the capability of the method to accommodate new information. Patients were divided into those newly diagnosed and those with recurrent disease, anticipating different MTDs in the two groups. We also incorporated data from patients who received $1000\,\mu g/m^2$ as part of within-patient dose escalations from the phase II trial. These individuals had no toxicity at $850\,\mu g/m^2$ and were escalated as part of the original study protocol.

After observing the results at $1260\,\mu g/m^2$, investigator opinion about the dose-response behavior at high doses changed. In step 3, additional information about one patient treated at $1740\,\mu g/m^2$ became available. After step 3, the estimated target dose changed by less than 10% and the dose finding was terminated with a recommendation to use $1776\,\mu g/m^2$ in subsequent studies. A graphical display of observed data and model fits is given in Fig. 2. It is interesting to compare the final estimate of d_{50} (=1883) with our guess from step 0 (=1600) to see how reasonable investigator intuition was. The initial guess was conservative but fairly accurate and performed well in guiding subsequent steps.

An assessment of CRM efficiency compared to traditional dose-finding methods can be obtained for this example. Suppose that $1776\,\mu g/m^2$ is the correct dose and that a series of doses with 15% increments had been employed instead, using a more traditional design. The doses used would have been 1150, 1323, 1521, 1749, and $2011\,\mu g/m^2$. Six steps would have been necessary; five would cover the range and an additional one after dose decreasing from 2011 to $1749\,\mu g/m^2$. The minimum number of patients required using this scheme would be $6 \times 3 = 18$ (plus any ineligibles and dropouts) compared with 11 actually used. Furthermore, this method would have slightly underestimated the correct dose and would have treated patients at $2011\,\mu g/m^2$, well above the MTD.

Even if 25% increments had been employed in a traditional design, more patients would have been required and the final dose chosen would have likely been

Table 2 Steps in the 9AC CRM for newly diagnosed patients

Step number	New dose	Data for CRM			Target dose	Model parameters	
		Dose	n	r		$\hat{\beta}$	$\hat{d}_{50}$
0		850	22	0			
		1600			1260	0.0025	1600
1		850	12	0			
		1000	7	0			
	⇒	1260	4	0			
		2000	10	9	1865	0.0209	1895
2		850	12	0			
		1000	7	0			
		1260	4	0			
	⇒	1865	3	2			
		2250	10	9	1740	0.0063	1848
3		850	12	0			
		1000	7	0			
		1260	4	0			
	⇒	1740	3	1			
		1865	4	2			
		2250	10	9	1776	0.0065	1883

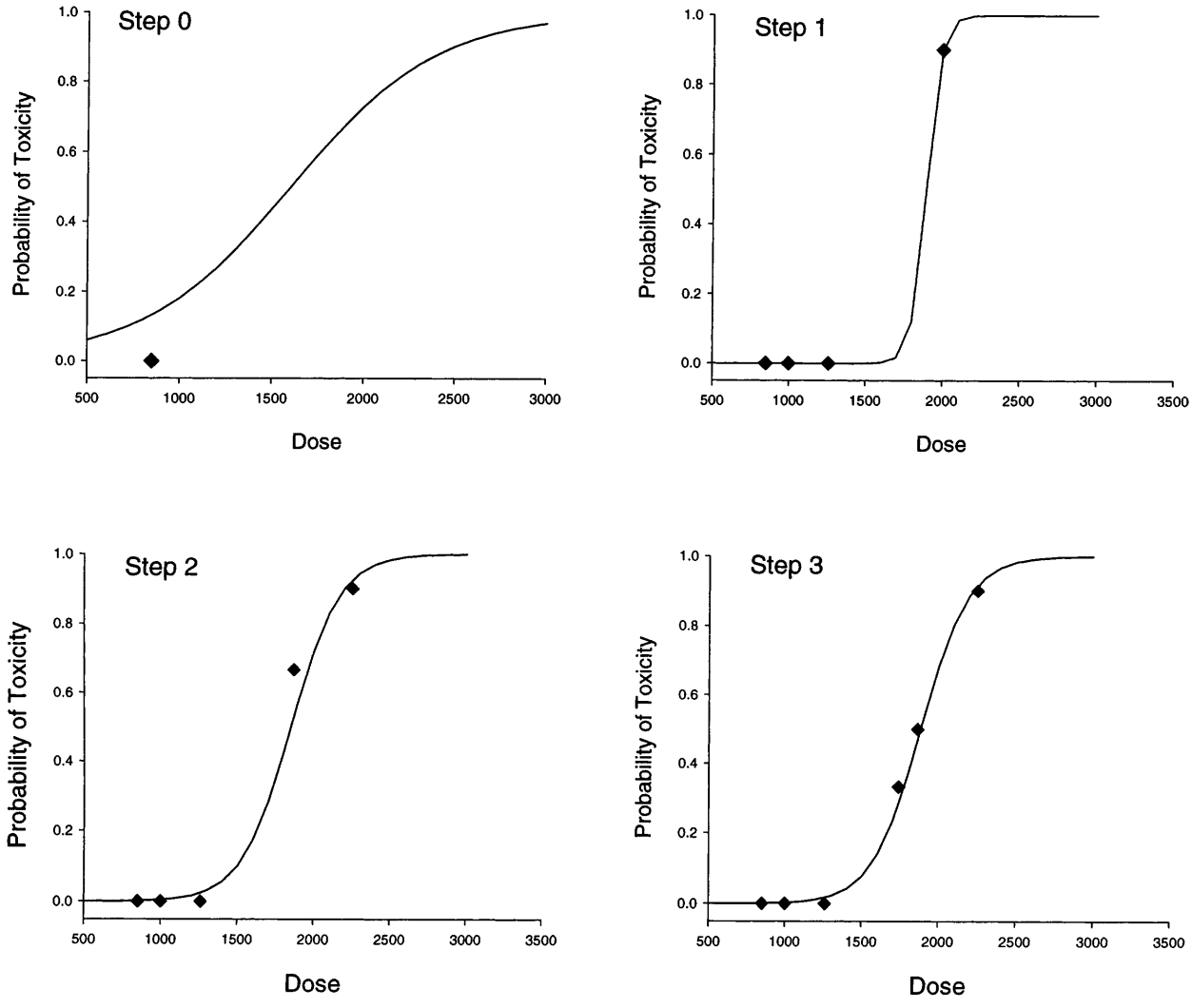

Fig. 2 Steps in the CRM for 9AC in newly diagnosed patients showing point estimates from data (◆) and fitted dose-response models

slightly biased. The usual Fibonacci dose escalation scheme would most likely not be used in this circumstance, because investigators could not have confidence that the starting dose was well below the MTD. Therefore, the smaller increments suggested here would probably have been used.

Patients with recurrent disease

The same initiation was used for dose escalations in patients with recurrent disease. At 1260 $\mu g/m^2$, one patient out of three experienced a DLT. Therefore, the same dose was repeated in the next cohort of three patients. No additional toxicities were observed and the CRM proceeded to 1611 $\mu g/m^2$ (Table 3). One patient was observed to have DLT at this dose and the estimation process appeared to converge. The recommended dose for subsequent investigations in patients with recurrent disease will be 1611 $\mu g/m^2$.

Discussion

Investigators need to solve several practical problems when implementing the CRM. An experienced statistician can help overcome these difficulties. First, the dose-response model must be initiated on the basis of relatively little data. This can be accomplished by ad hoc fitting of the model to the available data, as in the first example above, or by putting investigator knowledge in the form of data that can be used directly. Because the model has two parameters, at least two "data points" or bits of information are needed to initialize the model. Our experience is that the required information can be obtained from preliminary data or investigator knowledge.

Second, the dose-response model must be fitted to small amounts of data, especially in the early steps of the CRM. The model-fitting process will be ill-conditioned or impossible unless information at both ends of the curve is used. Making an informed guess about a dose that yields a high probability of response appears to be

Table 3 Steps in the 9AC CRM for patients with recurrent disease

Step Number	New dose	Data for CRM			Target dose	Model parameters	
		Dose	n	r		$\hat{\beta}$	$\hat{d}_{50}$
0		850	22	0			
		1600			1260	0.0025	1600
1		850	15	0			
		1000	6	0			
	⇒	1260	3	1	1260		
2		850	15	0			
		1000	6	0			
		1260	3	1			
	⇒	1260	3	0			
		2250	10	9	1611	0.0048	1757
3		850	15	0			
		1000	6	0			
		1260	6	1			
	⇒	1611	3	1			
		2250	10	9	1611	0.0048	1757

the best solution for the right-hand portion of the dose-response model. Investigators should feel free to revise this guess based on the observations from lower doses. An alternative fitting method is to restrict the parameter estimates and use a fitting technique that can accommodate unbalanced data (e.g. Bayesian methods). However, an approach that translates information into data seems more appealing to clinicians.

Third, the method must cope with a dose-response model that is most likely incorrect. Most imperfections in the model will be inconsequential. If it is approximately correct or locally valid, it will still guide and facilitate the dose escalations efficiently. Thus, it is more important to choose a simple, flexible model than to worry about its exact mathematical form. It would be very unlikely for enough data to became available during a dose-finding study to convince investigators that the model is inadequate.

Fourth, concerns have been raised that the CRM can escalate doses too quickly, increasing the chance for serious side effects [2]. Such problems can be minimized using groups of patients at each dose and empirically limiting the size of dose increases [3]. It appears perfectly adequate to use the assumed value for d_{90} as a convenient device for limiting dose jumps, as we have done. An explicit limitation of the magnitude of dose jumps could also be used. In addition to efficiency and accuracy, the CRM can be adapted to utilize pharmacokinetic information that may facilitate dose finding [12]. We have not yet explored this possibility in a clinical setting.

The CRM is operationally more complex than its predecessors. The principal complexities are:

1. It requires a mathematical model to summarize, interpolate, and extrapolate the data.
2. Investigators must translate biological/clinical information into data that can be used by the model.
3. It requires special methods of statistical estimation to optimally represent the data. This usually requires computer programs.
4. A statistical collaborator is needed to specify the design and perform calculations after each dose. In other words, the new dose is not evident by inspection.

These operational difficulties are minimized by the modifications we suggest, making the CRM an efficient, flexible, and practical method for dose finding in oncology. We believe it should be the preferred method for such studies.

Appendix I

The log-likelihood function for binomial outcomes and logistic dose response is

$$\mathscr{L}(\beta, d_{50}) = \sum_{i=1}^{k} \left[\log \binom{n_i}{r_i} - n_i \log\left(1 + e^{-\beta(d_i - d_{50})}\right) - (n_i - r_i)\ \log\left(1 + e^{\beta(d_i - d_{50})}\right) \right] \tag{A1}$$

This indicates that k doses have been tried, $d_1, d_2, \ldots, d_k$, with r_i responses out of n_i patients at each dose. The likelihood is the probability of the observed data under the hypothesized model. Values of β and d_{50} are chosen that maximize $\mathscr{L}$. These so-called maximum likelihood estimates, $\hat{\beta}$ and $\hat{d}_{50}$, must usually be determined numerically rather than algebraically. This method is general and is used during the dose escalations (Appendix II).

For example, suppose the data consist of the following triplets for $d_i,\ n_i,\ r_i$: $(10, 3, 0), (25, 3, 0), (50, 3, 1)$, and $(200, 10, 9)$. The maximum likelihood estimates are $\hat{\beta} = 0.027$ and $\hat{d}_{50} = 113$. The fitted curve and the data points are shown in Fig. 3. The best estimate of the dose that yields a 0.3 probability of response is 81.6.

When initiating the CRM with clinician-generated data for d_{10} and d_{90} (and no other data), $\hat{\beta}$ and $\hat{d}_{50}$ can be obtained directly. Because there are two data points and two parameters, the model will fit "perfectly". In other words,

$$\log(9) = \hat{\beta}(d_{90} - \hat{d}_{50}) \tag{A2}$$

and

$$-\log(9) = \hat{\beta}(d_{10} - \hat{d}_{50}). \tag{A3}$$

Solving Eq. A2 for β and substituting into Eq. A3, we obtain

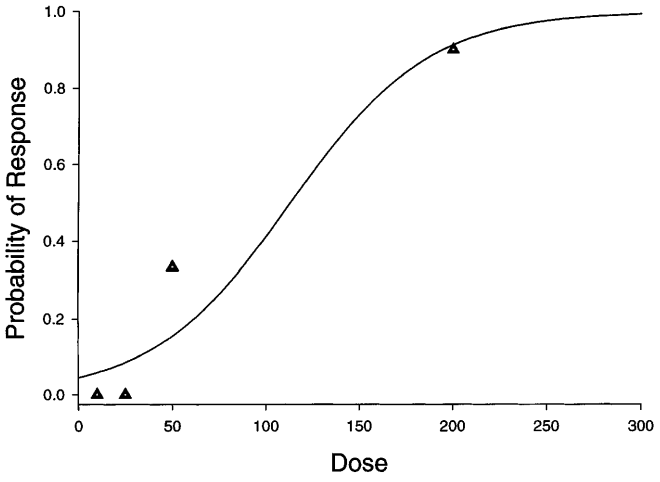

Fig. 3 Example fit of dose-response model to data points

$$\hat{d}_{50} = \frac{d_{90} + d_{10}}{2}. \quad \text{(A4)}$$

Thus, the initial best estimate of d_{50} is halfway between d_{10} and d_{90}. Substituting this into either Eq. A2 or Eq. A3,

$$\hat{\beta} = \frac{2\log(9)}{d_{90} - d_{10}}. \quad \text{(A5)}$$

The values of $\hat{d}_{50}$ and $\hat{\beta}$ can be used to obtain an initial estimate of the target. For example, if the target is the dose that yields a 0.3 probability of response, Eq. 2 can be used with Eqs. A4 and A5 to provide the first estimate. This first estimate of the target depends only on the values chosen for d_{10} and d_{90} and not on the strength of evidence (i.e. the "sample sizes"). We have implicitly encouraged choosing a first dose near the target. This is not required, although it will increase the efficiency of the dose escalation. Furthermore, the investigator is free to revise the parameters to produce a subjectively good starting dose.

Appendix II – CRM program

The following program executes under *Mathematica* versions 2.2.1 or 3.0 for Windows [16] and performs the CRM calculations. An actual program file is available from the authors on request. The algorithm is relatively simple and can be programmed in a variety of languages, such as Fortran or C/C++. However, Mathematica has a built-in multivariable function minimization algorithm (as do some other high-level languages), which makes the program and development time very short compared to low-level languages.

The model is parameterized in terms of $\log(d_{50})$ and $\log(\beta)$. The fitting process is sensitive to initial guesses for these parameter values, especially when few data are available. Data are given as a three-column matrix {dose, patients, responses} of arbitrary length.

Dose response model and likelihood function

```
p[dose_] := 1/(1 + Exp[- Exp [beta] * (dose - Exp[d50])])
ip[p_] := Exp[d50] - Log[(1 - p)/p]/Exp[beta]
term[x_] := Log[Binomial[data[[x, 2]], data[[x, 3]]]] + data[[x, 3]]
  *Log[p[data[[x, 1]]]] + data[[x, 2]] - data[[x, 3]])
  *Log[1 - p[data[[x, 1]]]]
loglike[d50_,beta_] :=Sum[term[i], i, 1, Dimensions[data][[1]]]
```

Constants and starting values

```
startd50 = 7.5;      startbeta = -5;      target = 0.3333;
highdose = 2250;     num = 9;     den = 10;     lowplot = 500;
highplot = 3000.
```

Data and results

```
data = {{850, 15, 0}, {1000, 6, 0}, {1260, 6, 1}, {1611, 3, 2},
{highdose, den, num}};  Clear [d50, beta]; q = FindMinimum
[-loglike[d50, beta], {d50, startd50}, {beta, startbeta}];
d50 = d50/.q[[2, 1]];  beta = beta/.q[[2, 2]];  MatrixForm[data];
ip[target];  p1 = Plot[p[dose], {dose, lowplot, highplot}].
```

References

1. 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
2. Korn EL, Midthune D, Chen TT, Rubinstein LV, Christian MC, Simon RM (1994) A comparison of two phase I trial designs. Stat Med 13: 1799–1806
3. Goodman SN, Zahurak ML, Piantadosi S (1995) Some practical improvements in the continual reassessment method for phase I studies. Stat Med 14: 1149–1161
4. Dixon WJ, Mood AM (1948) A method for obtaining and analysing sensitivity data. J Am Stat Assoc 43: 109–126
5. Wetherill GB (1963) Sequential estimation of quantal response curves (with discussion). J R Soc Stat B 25: 1–48
6. Storer BE (1989) Design and analysis of phase I clinical trials. Biometrics 45: 925–937
7. Robbins H, Monro S (1951) A stochastic approximation method. Ann Math Stat 22: 400–407
8. Wong WK, Lachenbruch PA (1996) Designing studies for dose response. Stat Med 15: 343–359
9. Ruberg SJ (1995) Dose response studies. I. Some design considerations. J Biopharm Stat 5: 1–14
10. Ruberg SJ (1995) Dose response studies. II. Analysis and interpretation. J Biopharm Stat 5: 15–42
11. Ratain MJ, Mick R, Schilsky RL, Siegler M (1993) Statistical and ethical issues in the design and conduct of phase I and II clinical trials of new anticancer agents. J Natl Cancer Inst 85: 1637–1643
12. Piantadosi S, Liu G (1996) Improved designs for phase I studies using pharmacokinetic measurements. Stat Med 15: 1605–1618
13. Piantadosi S (1997) Clinical trials: a methodologic perspective. John Wiley & Sons, New York
14. Rubin E, Wood V, Bharti A, et al (1994) A phase I trial of 9-aminocamptothecin (9-AC). Proc Am Assoc Cancer Res 35: 245
15. Grossman SA, Hochberg F, Fisher J, Chen TL, Kim L, Gregory R, Grochow LB, Piantadosi S, for the NABTT CNS Consortium (1997) Increased 9-aminocamptothecin dose requirements in patients on anticonvulsants. Proc Am Soc Clin Oncol 16: 389a
16. Wolfram S (1996) The mathematica book, 3rd edn. Wolfram Media and Cambridge University Press, Cambridge