Journal of the Royal Statistical Society
SERIES C Applied Statistics

Appl. Statist. (2009)
58, *Part* 2, *pp.* 211–224

Bayesian dose finding in oncology for drug combinations by copula regression

Guosheng Yin and Ying Yuan

University of Texas M. D. Anderson Cancer Center, Houston, USA

[Received September 2007. Final revision August 2008]

Summary. Treating patients with a combination of agents is becoming commonplace in cancer clinical trials, with biochemical synergism often the primary focus. In a typical drug combination trial, the toxicity profile of each individual drug has already been thoroughly studied in single-agent trials, which naturally offers rich prior information. We propose a Bayesian adaptive design for dose finding that is based on a copula-type model to account for the synergistic effect of two or more drugs in combination. To search for the maximum tolerated dose combination, we continuously update the posterior estimates for the toxicity probabilities of the combined doses. By reordering the dose toxicities in the two-dimensional probability space, we adaptively assign each new cohort of patients to the most appropriate dose. Dose escalation, de-escalation or staying at the same doses is determined by comparing the posterior estimates of the probabilities of toxicity of combined doses and the prespecified toxicity target. We conduct extensive simulation studies to examine the operating characteristics of the design and illustrate the proposed method under various practical scenarios.

Keywords: Adaptive design; Bayesian inference; Combining drugs; Continual reassessment method; Copula model; Maximum tolerated dose; Phase I trial; Toxicity probability

1. Introduction

In conventional phase I clinical trials, the primary objective is often to find the maximum tolerated dose (MTD) of the drug under study, i.e. the dose with a probability of toxicity that is closest to the physician's prespecified target. In such an early phase of a drug study, relatively little is known about the appropriate dose level; hence, a sequence of doses is screened to find the target dose that is associated with the maximum level of tolerable toxicity. Towards this goal, various statistical methods have been proposed for phase I clinical trial designs (see Storer (1989), O'Quigley *et al.* (1990, 2002), Korn *et al.* (1994), Møller (1995), Goodman *et al.* (1995), O'Quigley and Shen (1996) and Mukhopadhyay (2000), among others). The commonly used continual reassessment method (CRM) assumes a parametric link function, such as a hyperbolic tangent or a power function, between the true dose toxicity probabilities and the prespecified toxicity probabilities (O'Quigley *et al.*, 1990). The toxicity probability curve can be efficiently adjusted by updating the posterior estimate of a single unknown parameter (O'Quigley and Shen, 1996). Other methods include dose escalation with overdose control (Babb *et al.*, 1998), the curve-free dose finding procedure based on the product beta prior (Gasparini and Eisele, 2000) which can be exactly characterized by an equivalent CRM design, the decision theoretic approaches (Whitehead and Brunier, 1995; Leung and Wang, 2002) and the random-walk rule and the biased coin design under isotonic regression (Durham *et al.*, 1997; Stylianou and

Address for correspondence: Guosheng Yin, Department of Biostatistics, University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, USA.
E-mail: gsyin@mdanderson.org

 0035–9254/09/58211

Flournoy, 2002). Comprehensive reviews for phase I trial designs were given in Edler (2001) and Rosenberger and Haines (2002). More up-to-date and extensive discussions on statistical approaches for dose finding can be found in Chevret (2006).

However, all the aforementioned methods are developed for single-agent dose finding trials. Given the enormous advances in medicine and large numbers of new drugs to be tested, interest in finding combinations of drugs for patient treatment has grown. The goal of combination therapy is to achieve better patient response, particularly for cancer patients who are refractory to conventional therapies. In oncology, for example, combining agents can induce a synergistic treatment effect, allowing the clinician to target tumour cells with differing drug susceptibilities, and to achieve a higher intensity of dose with non-overlapping toxicities. Our research is motivated by many recent and more emerging dose finding drug combination clinical trials at the M. D. Anderson Cancer Center. One example is to find the MTD combination of a small molecule receptor (orally administered with four dose levels) and a mammalian target of rapamycin inhibitor (an intravenous drug with four dose levels) resulting in 16 combinations. The combined drugs are expected to induce a synergistic treatment effect by targeting different pathways.

The trend of drug combination trials poses a great challenge to finding the MTD combination with two or more drugs, particularly with small sample sizes in phase I studies. In a single-agent trial, we typically assume a monotonically increasing order of toxicity with respect to the dose. For any given dose, there are at most two adjacent doses and the order of toxicity is known. In contrast, for a two-drug combination dose space, there are up to eight adjacent doses, including diagonal and off-diagonal doses, as shown in Fig. 1. More importantly, complex drug–drug interactive effects often lead to unknown patterns of toxicity. Thus, the monotonic order of toxicity with respect to the dose level is lost, and it becomes unclear which dose combination should be assigned under a decision of dose escalation or de-escalation. Moreover, when two or more drugs are combined, the dimension of the dose space expands in a multiplicative fashion. This rapid increase in the dose dimension naturally requires a larger sample size, which can easily double or triple that of a single-agent trial.

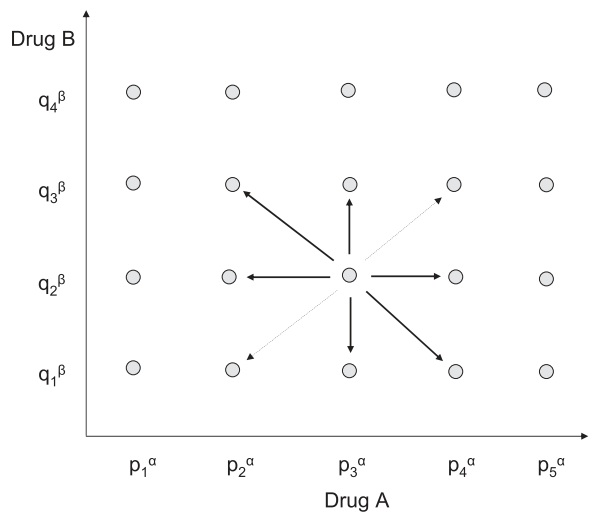

Fig. 1. Dose escalation and de-escalation diagram with 5 × 4 combinations, while the dose change along the diagonal (· · · ·►) is not allowed

A straightforward way to extend the traditional single-agent dose finding methods to drug combination trials is to conduct a series of one-dimensional dose finding trials: fixing one drug at each specified dose level and varying the other. This essentially transforms the two-dimensional dose finding space to a one-dimensional space. Korn and Simon (1993) introduced a graphical tolerable dose diagram to provide guidance in targeting specific MTD combinations. Kramar *et al.* (1999) proposed monotonically ordering a selected subset of drug combinations which reduced the dose finding to a one-dimensional space. Lokich (2001) presented a clinical trial that combined topotecan at four dose levels with irinotecan at two levels in the treatment of patients with advanced malignancy. Kuzuya *et al.* (2001) proposed combining paclitaxel and carboplatin to treat ovarian cancer, by alternately fixing one agent at a certain dose level and varying the other. Thall *et al.* (2003) proposed a six-parameter model to define the probability of toxicity with respect to the doses of gemcitabine and cyclophosphamide. The combined dose is first escalated along the diagonal line; then a toxicity equivalence contour is constructed for further investigation of proper dose combinations. Conaway *et al.* (2004) examined the simple and partial orders for drug combinations based on the pool adjacent violators algorithm. Wang and Ivanova (2005) proposed a two-dimensional dose finding design that was based on the logistic-type regression with standardized doses of the two drugs as the covariates. In this design, the interaction term can be included to capture the drug–drug interactive effects, which, however, might not be estimated reliably owing to a small variation in the covariate. Huang *et al.* (2007) extended the traditional '$3 + 3$' design by classifying the two-dimensional dose space into zones. Yin and Yuan (2008) proposed a latent 2×2 table approach to dose finding for drug combinations.

Trial designs involving drug combinations can be extended to a broader spectrum, for example, to search jointly for the optimal dose level and dose schedule for a single agent. As seen frequently in clinical trials, a single agent with a set of dose levels may be administered at several different dose schedules in a treatment cycle. To improve the treatment effect, the drug can be given at a more intense or frequent schedule instead of a regular schedule. More recently, Braun *et al.* (2007) proposed a Bayesian adaptive design that simultaneously optimizes both dose and schedule based on time-to-toxicity outcomes. Motivated by a phase I trial in allogeneic bone marrow transplantation at the M. D. Anderson Cancer Center, their design reflects the actual clinical practice and thus has potential of broad applications. Therefore, the drug combination can be formulated in a more general framework, that of a trial with several different drugs, each at a prespecified set of dose levels, or a study of a single agent at a set of dose levels, adding a change to different dose schedules. The terminology of 'dose level' or 'dose schedule' may be used exchangeably for our general purpose of describing statistical designs, as both would increase the probability of toxicity monotonically. Other work related to drug combination studies includes Plackett and Hewlett (1967), Ashford (1981), Abdelbasit and Plackett (1982), Simon and Korn (1990) and Korn and Simon (1992).

Before several drugs are combined, the toxicity profile of each individual drug needs to be thoroughly investigated. Hence, physicians usually have good prior knowledge of the probabilities of toxicity when each drug is administered alone. We propose to utilize this rich prior information and to develop a Bayesian adaptive dose finding procedure for combined agents. We model the rates of toxicity of the combined drugs via a copula-type regression. In particular, we first prespecify the underlying probabilities of toxicity for each drug on the basis of their individual toxicity information. To accommodate the uncertainty of this prespecification, we incorporate a power parameter to the probabilities of toxicity for each drug as in the CRM. We then link these probabilities of toxicity in a copula-type model to derive the joint toxicity probability of the combined drugs. Our method reduces to the CRM if only one drug is

tested. Patients are sequentially accrued in cohorts. Decisions on dose escalation, de-escalation or staying at the same dose are adaptively made as the data are coherently updated in the trial.

The rest of the paper is organized as follows. In Section 2, we propose the joint toxicity probability model for the binary outcomes through a copula-type regression. Moreover, we derive the likelihood function and the posterior distribution for the unknown parameters. We formulate the decision rules and the dose finding algorithm in Section 3. We conduct extensive simulation studies to examine the operating characteristics of the new design in Section 4, and we conclude with a brief discussion in Section 5.

The programs that were used to simulate the design can be obtained from

`http://www.blackwellpublishing.com/rss`.

The executable file can be downloaded from `http://odin.mdacc.tmc.edu/~gyin/software.htm`.

2. Probability model

2.1. Copula-type regression

For ease of exposition, we illustrate our design by using a drug combination trial with two agents, say A and B, though our approach can be easily generalized to the case of combining multiple agents. Let p_j be the prespecified probability of toxicity corresponding to A_j, the jth dose for drug A, $p_1 < p_2 < \ldots < p_J$, and q_k be that of B_k, the kth dose for drug B, $q_1 < q_2 < \ldots < q_K$. Typically, the maximum dose for each drug in the combination (i.e. A_J and B_K) is the individual MTD that has already been determined in the single-agent trials. The remaining lower doses are some fractions of the MTD, specified by physicians. The probability of toxicity is monotonically increasing with the dose level when each drug is administered alone. However, the ordering of the probabilities of toxicity of the combined drugs and dose levels is less obvious. For example, it is not clear how to order the probabilities of toxicity π_{jk}, $\pi_{j-1,k+1}$ and $\pi_{j+1,k-1}$, where π_{jk} is the joint toxicity probability associated with the combined drug pair (A_j, B_k).

In a drug combination study, the probability of toxicity corresponding to the MTD of each agent is known, i.e. p_J and q_K are known. Through further consultation with physicians, we can easily specify $(p_1, \ldots, p_{J-1})$ for drug A, and $(q_1, \ldots, q_{K-1})$ for drug B. To enhance the flexibility and to accommodate physicians' uncertainty, we take p_j^α and q_k^β as the true probabilities of toxicity for drug A and drug B respectively, where $\alpha > 0$ and $\beta > 0$ are unknown parameters with prior means centred at 1. When two or more drugs are combined as a treatment, it would be unrealistic to assume that each drug acts independently on the patient, as the drug–drug interactive effects may have a strong influence on the joint toxicity profile. Therefore, it is critical to choose an appropriate model to link the joint toxicity probability π_{jk} with the p_js and q_ks. A reasonable model for drug combinations needs to satisfy the following conditions: for $j = 1, \ldots, J$ and $k = 1, \ldots, K$,

(a) if $p_j^\alpha = 0$ and $q_k^\beta = 0$, then $\pi_{jk} = 0$,
(b) if $p_j^\alpha = 0$, then $\pi_{jk} = q_k^\beta$, and, if $q_k^\beta = 0$, then $\pi_{jk} = p_j^\alpha$, and
(c) if either $p_j^\alpha = 1$ or $q_k^\beta = 1$, then $\pi_{jk} = 1$.

These conditions are very intuitive for the joint model of drug combinations. The first condition requires that, if the probabilities of toxicity of both drugs are 0 (i.e. no drugs), the joint toxicity probability should also be 0. The second condition states that, if the probability of toxicity of one drug is 0 (i.e. a single-agent case), the joint toxicity probability should reduce to that of the other drug. Finally, the third condition asserts that, if the toxicity probability of either drug is 1, the joint toxicity probability should be 1 as well.

Although each drug has its own toxicity profile, the individual toxicity outcome cannot be observed in the trial. We typically can only observe a single toxicity binary outcome for the drug combination. For example, the outcome of a patient treated at (A_j, B_k) provides information on the joint toxicity probability π_{jk} that can be linked to the probabilities of toxicity $(p_j^{\alpha}, q_k^{\beta})$ through a copula-type regression model. There has been considerable interest in modelling the bivariate probability function through copula models (for example, see Clayton (1978), Hougaard (1986), Genest and Rivest (1993) and Nelsen (1999), among others). Copula models have an attractive structure by expressing the joint probability distribution through the marginal distributions and a dependence parameter. Suppose that C_γ is a distribution function on $[0, 1]^2$ given an association parameter γ. The Archimedean copula family characterizes an important class of dependence functions, which has the representation

$$C_\gamma(u, v) = \psi_\gamma\{\psi_\gamma^{-1}(u) + \psi_\gamma^{-1}(v)\}, \qquad 0 \leqslant u, v \leqslant 1,$$

where ψ_γ is the copula generator, ψ_γ^{-1} is its inverse function, $0 \leqslant \psi_\gamma \leqslant 1$, $\psi_\gamma(0) = 1$ and first and second derivatives $\psi_\gamma' < 0$ and $\psi_\gamma'' > 0$. It encompasses many well-known bivariate parametric distributions, such as the Clayton, Gumbel–Hougaard and Frank copula models (see Nelsen (1999)).

Motivated by the Clayton copula, we propose a copula-type regression model to link the joint toxicity probability π_{jk} with $(p_j^{\alpha}, q_k^{\beta})$ in the form of

$$\pi_{jk} = 1 - \{(1 - p_j^{\alpha})^{-\gamma} + (1 - q_k^{\beta})^{-\gamma} - 1\}^{-1/\gamma}, \tag{1}$$

where $\gamma > 0$ characterizes the drug–drug interactive effect. $\lim_{p_j \to 1}\{(1 - p_j^{\alpha})^{-\gamma}\} = \infty$ and $\lim_{q_k \to 1}\{(1 - q_k^{\beta})^{-\gamma}\} = \infty$, and thus $\pi_{jk} = 1$ as p_j or q_k goes to 1. Moreover, if only one drug is tested, say $p_j > 0$ and $q_k = 0$ $(k = 1, \ldots, K)$, model (1) reduces to the CRM, with $\pi_j = p_j^{\alpha}$ $(j = 1, \ldots, J)$. Our approach can be regarded as a multivariate generalization of the CRM which enables internal learning from other dose level combinations. As a result, this model satisfies the three model conditions. In Fig. 2, we illustrate the joint toxicity probability surface based on model (1) with $\alpha = \beta = 2$ and $\gamma = 1.5$ in the two-dimensional probability space. Depending on the three parameters (α, β, γ), the toxicity probability surface may have various shapes. If the target toxicity probability is 40%, as shown by the horizontal plane in Fig. 2, there is an intersection curve representing the MTD contour for the two drugs. Therefore, depending on the physician-specified doses, there could be more than one MTD combination in the discrete dose space. We focus on selecting one MTD combination, whereas, if multiple MTDs are found, physicians can determine the one for further investigation on the basis of disease and subject-matters. Other copula models can be used as well, depending on mathematical convenience and computational simplicity. For example, on the basis of the Gumbel–Hougaard copula, we can model the joint toxicity probability by

$$\pi_{jk} = 1 - \exp(-[\{-\log(1 - p_j^{\alpha})\}^{1/\gamma} + \{-\log(1 - q_k^{\beta})\}^{1/\gamma}]^{\gamma}),$$

which clearly also satisfies the three model conditions.

2.2. Likelihood and prior specification

We can construct the likelihood function on the basis of the binomial distribution with the probabilities π_{jk}. Suppose that, at the current stage of the trial, among n_{jk} patients treated at the paired dose level (j, k), x_{jk} patients have experienced toxicity. The likelihood is then given by

$$\text{lik}(\alpha, \beta, \gamma | \text{data}) \propto \prod_{j=1}^{J} \prod_{k=1}^{K} \pi_{jk}^{x_{jk}} (1 - \pi_{jk})^{n_{jk} - x_{jk}}.$$

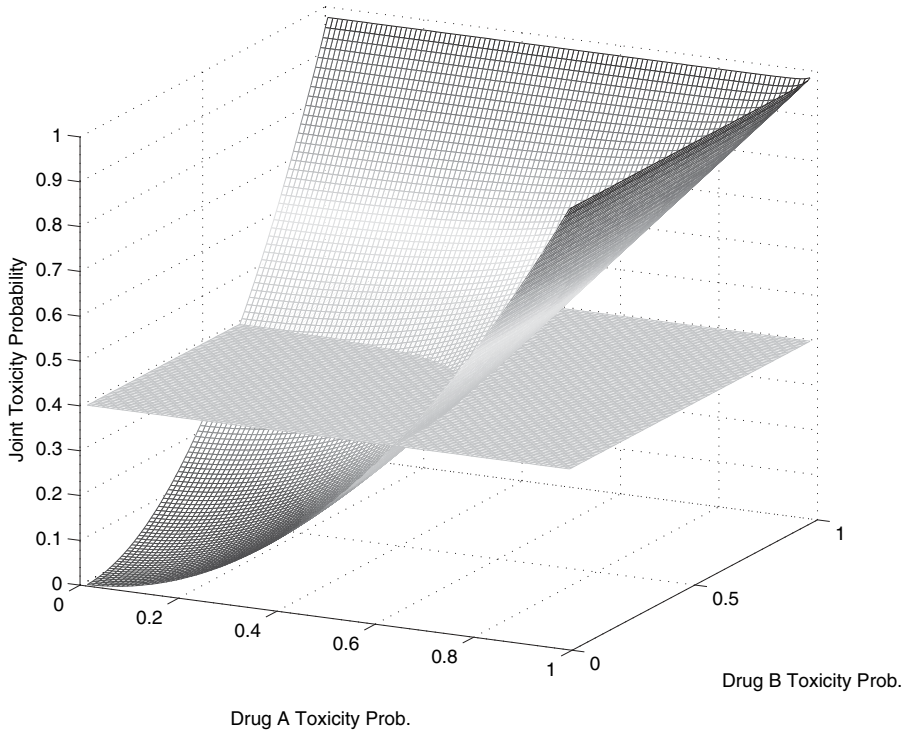

Fig. 2. Toxicity probability surface and MTD contour for the two-drug combination

For ease of computation, we take the prior distributions of the model parameters to be independent, i.e. $f(\alpha, \beta, \gamma) = f(\alpha)\, f(\beta)\, f(\gamma)$. Since the drugs in the combination have already been studied in single-agent trials, physicians often have some prior information of the probabilities of toxicity for each individual drug. We assign intermediate informative prior distributions to α and β by centring the prior mean at 1 with a relatively small variance, e.g. gamma(2, 2), a gamma distribution with mean 1 and variance 0.5. Essentially, these priors centre $(p_j^{\alpha}, q_k^{\beta})$ at the probabilities of toxicity when each drug is administered alone. The drug–drug interactive effect is mainly determined by γ, for which much less information can be elicited from physicians. Hence, we take a non-informative prior for γ, a gamma distribution with a large variance, so that the data will dominate the posterior estimation of γ. The joint posterior distribution is given by

$$f(\alpha, \beta, \gamma|\text{data}) \propto \text{lik}(\alpha, \beta, \gamma|\text{data})\, f(\alpha)\, f(\beta)\, f(\gamma),$$

from which the full conditional distributions of the model parameters can be easily obtained. After the outcomes of each cohort of patients have been observed, we use the Gibbs sampling algorithm to sample from the posterior distributions of the unknown parameters. Thus, we can easily obtain the posterior estimates for π_{jk}, on which the next stage of the trial design will be based.

2.3. Multiple drugs in combination

In practice, drug combination trials often involve a pair of drugs, each with several prespecified doses. However, clinicians may find it beneficial to combine more than two drugs to treat

a patient. When the dimension of the drug combination space is higher than 2, dose finding becomes a much more complicated process, for which most of the currently available methods might not work well. Our method, however, can be easily generalized to such a high dimensional dose combination problem. For example, if three drugs are combined in a trial, we denote p_j to be the physician-specified probability of toxicity for the jth dose of drug A, $j = 1, \ldots, J$, q_k to be that for the kth dose of drug B, $k = 1, \ldots, K$, and s_l to be that for the lth dose of drug C, $l = 1, \ldots, L$, i.e. the triplet (p_j, q_k, s_l) represents the prespecified probabilities of toxicity that are associated with the combined drug doses (A_j, B_k, C_l). By incorporating a power parameter for each prior probability of toxicity, the true probabilities of toxicity are $(p_j^{\alpha}, q_k^{\beta}, s_l^{\zeta})$, where α, β and ζ have gamma distributions with prior means centred at 1. The dose combination searching space grows multiplicatively into a three-dimensional cube of dimension $J \times K \times L$.

In this three-dimensional toxicity probability space, we can still quantify the joint toxicity probabilities through a copula-type model. Under the Clayton copula, we define

$$\pi_{jkl} = 1 - \{(1 - p_j^{\alpha})^{-\gamma} + (1 - q_k^{\beta})^{-\gamma} + (1 - s_l^{\zeta})^{-\gamma} - 2\}^{-1/\gamma}$$

and, under the Gumbel–Hougaard copula, we have

$$\pi_{jkl} = 1 - \exp(-[\{-\log(1 - p_j^{\alpha})\}^{1/\gamma} + \{-\log(1 - q_k^{\beta})\}^{1/\gamma} + \{-\log(1 - s_l^{\zeta})\}^{1/\gamma}]^{\gamma}),$$

where α, β and ζ are the unknown parameters for each individual drug and γ is the association parameter for synergism. The corresponding likelihood based on the binomial distribution is given by

$$\text{lik}(\alpha, \beta, \zeta, \gamma | \text{data}) \propto \prod_{j=1}^{J} \prod_{k=1}^{K} \prod_{l=1}^{L} \pi_{jkl}^{x_{jkl}} (1 - \pi_{jkl})^{n_{jkl} - x_{jkl}},$$

where, among n_{jkl} patients treated at (A_j, B_k, C_l), x_{jkl} patients have experienced toxicity. The rest of the prior specifications and posterior derivations can be carried forward similarly.

3. Dose finding algorithm

Let ϕ be the physician-specified targeting toxicity limit, and c_e and c_d be the fixed probability cut-offs for dose escalation and de-escalation respectively. c_e and c_d can be calibrated through simulation studies such that the trial has desirable operating characteristics, and $c_e + c_d > 1$. Patients are treated in cohorts, for example, with a cohort size of 3. To be conservative, we restrict dose escalation or de-escalation to one dose level of change only, while also not allowing a move along the diagonal direction (corresponding to simultaneous escalation or de-escalation of both agents), as shown in Fig. 1. For a trial involving two drugs, the dose finding algorithm works as follows.

(a) Patients in the first cohort are treated at the lowest dose combination (A_1, B_1).
(b) If, at the current dose combination (j, k),

$$\Pr(\pi_{jk} < \phi) > c_e,$$

the dose is escalated to an adjacent dose combination with the probability of toxicity higher than the current value and closest to ϕ. If the current dose combination is (A_J, B_K), the doses stay at the same levels.
(c) If, at the current dose combination (j, k),

$$\Pr(\pi_{jk} > \phi) > c_d,$$

the dose is de-escalated to an adjacent dose combination with the probability of toxicity lower than the current value and closest to ϕ. If the current dose combination is (A_1, B_1), the trial is terminated.

(d) Otherwise, the next cohort of patients continues to be treated at the current dose combination (doses staying at the same levels).

(e) Once the maximum sample size has been reached, the dose combination that has the probability of toxicity that is closest to ϕ is selected as the MTD combination.

As is common to model-based clinical trial designs, the dose finding algorithm is difficult to apply at the beginning of the trial, because very limited information except for the prior knowledge is available. Thus, the posterior estimates of the probabilities of toxicity for dose combinations may not be reliable. To circumvent this difficulty, we use the following start-up rule: we first treat patients along the vertical dose escalation in the order of $\{(A_1, B_1), (A_1, B_2), \dots\}$ until the first toxicity is observed; we then continue to treat patients by escalating doses in the horizontal direction $\{(A_2, B_1), (A_3, B_1), \dots\}$ until the first toxicity occurs. As long as one toxicity is observed in both the vertical and the horizontal directions, e.g., if one patient experiences toxicity at (A_1, B_k) and (A_j, B_1) for some values of j and k, the Bayesian dose finding algorithm will be in effect seamless for the rest of the trial.

4. Simulation study

We investigated the operating characteristics of our two-dimensional Bayesian copula dose finding method through simulation studies under 12 different toxicity scenarios, as listed in Table 1. In our simulation, the target probability of toxicity was $\phi = 40\%$, and the sample size of each trial was 60, with a cohort size of 3. We set $c_e = 0.8$ and $c_d = 0.45$ to direct dose escalation and de-escalation. The dose assignment decision was made after observing the outcomes of every cohort of patients. In the Clayton copula, we took gamma(2, 2) as the prior distribution for α and β, and gamma(0.1, 0.1) for γ. In the Markov chain Monte Carlo procedure, we recorded 2000 posterior samples of the model parameters after 100 burn-in iterations. We simulated 2000 trials under each scenario. The first six scenarios represent drug combinations with five dose levels of drug A and four dose levels of drug B, but with different numbers and different locations of the MTD combinations. The toxicity rates of the MTDs for drug A and drug B were 0.4 and 0.3 respectively, when administered alone. For drug A, we specified the prior marginal toxicity probabilities p_j as (0.08, 0.16, 0.24, 0.32, 0.4); and, for drug B, we specified q_k as (0.075, 0.15, 0.225, 0.3). In scenarios 7–10, we examined a different dose configuration with a 4×4 grid. The prior toxicity probabilities were (0.07, 0.15, 0.22, 0.3) for drug A, and (0.12, 0.18, 0.24, 0.3) for drug B. As a part of the sensitivity analysis, scenarios 11 and 12 were designed to examine the robustness of the Bayesian copula design to model misspecifications. Under these two scenarios the probabilities of toxicity were generated according to the Frank and Gumbel copulas (Nelsen, 1999). We compared the performance of our method with the CRM and non-parametric optimal design (O'Quigley *et al.*, 1990, 2002). We applied the CRM in a series of one-dimensional dose finding trials by fixing drug B at each given dose and searching over the doses of drug A (referred to as the restricted CRM). The total 60 patients were equally allocated to the four independent trials, 15 patients for each trial. We adopted a stopping rule: if $\Pr\{\text{toxicity rate at } (A_1, B_1) > \phi\} > 0.9$, the trial is terminated for safety (Goodman *et al.*, 1995; Møller, 1995). The non-parametric optimal design offers a theoretical bound as a benchmark for numerical comparisons.

In Table 2, we report the selection probabilities of dose combinations based on the restricted CRM, the non-parametric optimal design and the Bayesian copula method proposed. Scenar-

Table 1. 12 toxicity scenarios for the two-drug combinations with the target toxicity probability 0.4†

Dose level of drug B	*Toxicity probabilities for the following levels of drug A:*									
	1	*2*	*3*	*4*	*5*	*1*	*2*	*3*	*4*	*5*
	Scenario 1					*Scenario 2*				
4	0.54	0.67	0.75	0.81	0.86	0.49	0.58	0.68	0.75	0.81
3	0.48	0.59	0.68	0.75	0.81	*0.40*	0.49	0.59	0.68	0.75
2	*0.40*	0.45	0.59	0.67	0.74	0.27	*0.40*	0.45	0.59	0.67
1	0.24	*0.40*	0.47	0.56	0.64	0.18	0.29	*0.40*	0.47	0.56
	Scenario 3					*Scenario 4*				
4	0.31	*0.40*	0.50	0.61	0.75	*0.40*	0.52	0.72	0.75	0.84
3	0.23	0.34	*0.40*	0.53	0.67	0.29	*0.40*	0.51	0.60	0.68
2	0.16	0.25	0.34	*0.40*	0.52	0.20	0.31	*0.40*	0.50	0.59
1	0.09	0.16	0.18	0.22	*0.40*	0.12	0.21	0.30	*0.40*	0.47
	Scenario 5					*Scenario 6*				
4	0.11	0.16	0.23	0.31	*0.40*	0.79	0.85	0.89	0.92	0.94
3	0.06	0.11	0.17	0.24	0.32	0.74	0.81	0.86	0.89	0.92
2	0.03	0.07	0.11	0.17	0.24	0.67	0.76	0.81	0.86	0.89
1	0.01	0.03	0.07	0.11	0.18	0.57	0.67	0.74	0.79	0.84
	Scenario 7					*Scenario 8*				
4	0.28	0.30	0.32	*0.40*		0.52	0.69	0.70	0.75	
3	0.22	0.24	0.27	*0.39*		0.30	0.62	0.68	0.71	
2	0.20	0.22	0.26	0.28		0.20	*0.40*	0.50	0.66	
1	0.14	0.20	0.21	0.24		0.15	0.30	*0.40*	0.55	
	Scenario 9					*Scenario 10*				
4	0.50	0.66	0.67	0.73		0.50	0.52	0.57	0.63	
3	*0.40*	0.54	0.62	0.68		0.28	*0.40*	0.55	0.59	
2	0.21	*0.40*	0.50	0.60		0.26	0.30	*0.40*	0.50	
1	0.15	0.25	*0.40*	0.56		0.20	0.23	0.24	*0.40*	
	Scenario 11					*Scenario 12*				
4	0.46	0.53	0.58	0.61	0.64	0.46	0.53	0.60	0.65	0.70
3	0.45	0.50	0.55	0.58	0.61	*0.39*	0.47	0.54	0.60	0.66
2	*0.40*	0.44	0.50	0.53	0.56	0.31	*0.40*	0.47	0.54	0.61
1	0.20	*0.40*	0.47	0.48	0.50	0.23	0.32	*0.40*	0.47	0.54

†The MTD combinations are in italics.

ios 1, 2 and 3 describe situations in which there are two, three and four MTD combinations respectively. Under the first three scenarios, our method selected the MTD combinations with higher probabilities than the restricted CRM which tended to select over-MTD combinations in scenarios 1 and 2, and under-MTD combinations in scenario 3. Under the restricted CRM design, the trials corresponding to each dose level of drug B were completely independent of each other and no information or strength was borrowed across them. In contrast, the Bayesian copula dose finding method continuously updated the posterior estimates of probabilities of toxicity for all dose combinations on the basis of the available data and efficiently searched for the target across the whole drug combination space. Even in the cases where the set of doses of drug A contained the MTD for each fixed dose level of drug B, as illustrated by scenarios 3 and 4, our method performed slightly better than the restricted CRM, and comparably with the non-parametric optimal design. When the MTD is in the upper right-hand corner of the grid as in scenario 5, the Bayesian copula design proposed significantly outperformed the restricted

Table 2. Selection probabilities of the restricted CRM, non-parametric optimal design and Bayesian copula design under the first 10 scenarios

Scenario	*Results for the restricted CRM*					*Results for the non-parametric optimal design*					*Results for the Bayesian copula design*				
1	15.5	2.0	0.0	0.0	0.0	3.2	0.0	0.0	0.0	0.0	0.9	0.1	0.0	0.0	0.0
	15.5	4.8	0.5	0.0	0.0	6.9	0.1	0.0	0.0	0.0	10.2	1.7	0.1	0.0	0.0
	11.8	8.8	2.3	0.3	0.0	*29.1*	17.2	0.1	0.0	0.0	*24.9*	11.6	1.7	0.1	0.0
	5.0	*12.3*	6.0	1.5	0.3	6.0	*29.0*	7.6	0.7	0.0	3.1	*19.1*	6.2	0.3	0.0
2	15.3	4.5	0.3	0.0	0.0	3.8	0.5	0.0	0.0	0.0	5.0	1.5	0.0	0.0	0.0
	12.8	8.3	2.0	0.3	0.0	*16.0*	4.1	0.2	0.0	0.0	*17.6*	11.6	1.4	0.1	0.0
	5.8	*11.8*	6.0	1.3	0.3	2.9	*16.3*	17.1	0.1	0.0	7.3	*19.2*	9.3	1.7	0.1
	1.5	9.0	*9.8*	3.5	1.3	0.1	11.2	*16.8*	8.7	2.1	0.1	5.1	*11.2*	5.0	0.3
3	7.0	*11.3*	5.3	1.0	0.0	8.4	*12.3*	15.8	0.5	0.0	3.8	*11.7*	8.5	1.3	0.1
	3.3	10.5	*7.8*	2.8	0.8	0.3	9.2	*11.7*	2.5	0.0	0.7	5.9	*12.7*	8.0	1.2
	0.8	6.8	10.0	*5.0*	2.5	0.0	1.5	9.2	*12.1*	3.7	0.0	0.7	3.5	*12.0*	10.3
	0.0	1.8	5.0	7.3	*11.0*	0.0	0.0	0.0	0.1	*12.7*	0.0	0.0	0.3	3.5	*15.8*
4	*13.3*	9.0	0.8	0.0	0.0	*12.6*	2.2	0.0	0.0	0.0	*10.2*	8.7	0.9	0.0	0.0
	6.0	*12.3*	5.5	1.0	0.3	3.6	*13.1*	2.1	0.3	0.0	5.3	*16.7*	5.9	0.5	0.0
	2.0	10.0	*8.8*	3.5	1.0	0.0	11.0	*12.9*	4.9	0.5	0.4	4.7	*14.1*	8.3	1.0
	0.3	4.3	9.5	*7.0*	4.0	0.0	0.2	5.0	*12.6*	19.0	0.1	0.2	5.8	*11.5*	5.3
5	0.3	3.0	7.3	7.3	*7.8*	0.0	0.0	0.1	8.1	*77.2*	0.0	0.0	0.4	3.2	*87.6*
	0.0	0.8	4.3	7.0	13.0	0.0	0.0	0.0	0.4	13.5	0.0	0.0	0.0	0.4	7.5
	0.0	0.3	1.5	4.8	18.8	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.7
	0.0	0.0	0.3	1.8	23.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1
6	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	4.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	8.8	0.3	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	14.0	1.8	0.3	0.0	0.0	99.9	0.1	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0
7	4.3	8.3	7.0	*5.3*		0.9	4.6	13.7	*44.7*		0.4	4.3	7.0	*45.9*	
	2.0	6.5	8.0	*8.5*		0.0	0.3	0.6	*33.7*		0.2	1.8	3.0	*24.0*	
	1.3	6.0	7.5	10.3		0.0	0.0	0.7	0.6		0.0	0.3	1.1	9.0	
	0.5	3.8	6.3	14.5		0.0	0.0	0.0	0.1		0.0	0.1	0.2	1.8	
8	16.0	2.5	0.3	0.0		4.2	0.0	0.0	0.0		11.8	0.2	0.0	0.0	
	12.8	10.8	0.8	0.0		8.7	0.2	0.0	0.0		10.9	12.2	0.3	0.0	
	3.3	*13.3*	6.8	1.5		0.1	*30.1*	15.5	0.0		1.4	*26.2*	9.2	0.4	
	1.0	9.0	*10.0*	5.0		0.0	8.6	*30.7*	1.8		0.0	7.0	*15.0*	3.4	
9	16.5	3.0	0.3	0.0		9.4	0.0	0.0	0.0		5.1	0.7	0.1	0.0	
	13.5	8.3	1.3	0.3		*21.9*	2.4	0.0	0.0		*15.7*	10.2	1.0	0.1	
	3.5	*13.3*	6.5	1.8		0.2	*23.9*	9.9	0.2		1.9	*24.1*	10.7	0.5	
	0.8	7.5	*11.0*	5.8		0.0	7.5	*23.9*	0.7		0.0	5.9	*17.7*	4.2	
10	14.0	5.0	1.0	0.3		9.2	3.3	0.4	0.0		6.9	5.9	0.9	0.1	
	5.8	*13.3*	4.8	1.0		2.7	*20.9*	0.8	0.3		4.9	*16.9*	6.7	2.4	
	3.5	9.8	*8.0*	3.5		0.9	13.0	*20.4*	8.7		0.7	6.8	*14.3*	10.1	
	1.5	5.8	8.0	*9.8*		0.0	0.2	0.3	*19.0*		0.1	1.5	4.5	*13.3*	

CRM. Scenario 6 examines the situation that all drug combinations are over toxic. Our method successfully early terminated 99.9% of the simulated trials without any selection. This scenario again demonstrates the enormous advantage of jointly modelling the probabilities of toxicity across the two-dimensional space. Note that the non-parametric optimal design does not have a stopping rule. Scenarios 7–10 investigate the performance of our method under a smaller drug combination space with four dose levels for each drug. The main differences between the scenarios are the numbers of MTD combinations and their locations in the 4×4 grid. When two MTD combinations lie together in the upper right-hand corner as in scenario 7, both can be selected with a high percentage. Scenarios 8 and 10 are more difficult situations, in which the MTD contour is not complete, as the MTD is skipped in the first column (fixing drug A

at dose level 1) owing to the discreteness of the dose space. Yet, our method still selected the MTD combinations with high percentages. In scenario 9, three MTD combinations exist, and our method correctly selected them almost up to 60%. The overall performance of the Bayesian copula design is satisfactory and is better than the restricted CRM. In many cases, the performance of the Bayesian copula design is also comparable with the non-parametric optimal design.

Table 3 displays the average number of patients treated at each dose combination across 2000 simulations. Our method generally performed better than the restricted CRM in the sense that a higher percentage of patients were treated at the MTD or the dose combinations nearby. This advantage is especially prominent in scenarios 1, 5, 7 and 8, where the set of prespecified doses of drug A may not contain the MTD for a fixed level of drug B. For example, in scenario

Table 3. Number of patients treated at each dose combination for the restricted CRM and Bayesian copula design under the first 10 scenarios

Scenario	*Results for the restricted CRM*					*Results for the Bayesian copula design*				
1	10.5	2.4	0.2	0.0	0.0	0.8	0.1	0.0	0.0	0.0
	10.1	3.4	0.4	0.0	0.0	4.7	1.0	0.1	0.0	0.0
	8.5	4.7	1.2	0.1	0.0	*14.2*	5.2	1.0	0.1	0.0
	5.7	*6.3*	2.5	0.4	0.1	14.5	*9.2*	2.9	0.6	0.1
2	10.1	3.2	0.4	0.0	0.0	2.7	0.8	0.1	0.0	0.0
	8.6	4.7	1.1	0.1	0.0	*8.4*	3.6	0.6	0.1	0.0
	6.2	*6.0*	2.3	0.4	0.1	9.4	*7.9*	3.5	0.7	0.1
	4.3	5.7	*3.7*	1.1	0.2	8.5	6.0	*5.7*	2.5	0.7
3	6.6	*5.8*	2.2	0.3	0.0	3.7	*3.9*	2.8	0.9	0.3
	5.2	5.9	*3.1*	0.7	0.1	3.7	3.1	*4.1*	2.3	1.3
	4.0	5.3	4.1	*1.4*	0.3	3.7	1.3	2.3	*4.1*	4.0
	3.2	3.9	4.1	2.6	*1.2*	6.2	2.8	2.8	3.7	*6.0*
4	*8.7*	4.9	0.8	0.0	0.0	*4.9*	2.8	0.6	0.1	0.0
	6.2	*6.2*	2.2	0.3	0.0	5.9	*5.7*	2.5	0.3	0.2
	4.7	5.8	*3.5*	0.9	0.2	4.7	3.9	*5.0*	2.4	0.8
	3.5	4.7	4.4	*1.9*	0.5	6.6	3.3	4.8	*5.1*	3.1
5	3.4	4.1	4.2	2.4	*0.9*	2.8	0.6	0.7	2.2	*22.7*
	3.1	3.6	4.0	2.8	1.5	2.8	0.0	0.1	0.3	5.2
	3.0	3.2	3.6	3.1	2.1	3.0	0.0	0.0	0.0	3.1
	3.0	3.0	3.2	3.2	2.6	6.0	2.9	3.0	3.2	4.3
6	7.7	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	8.5	0.6	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0
	9.4	1.2	0.0	0.0	0.0	3.6	0.0	0.0	0.0	0.0
	10.3	2.1	0.2	0.0	0.0	12.3	0.4	0.0	0.0	0.0
7	5.7	5.2	3.0	*1.0*		1.8	2.7	3.1	*12.1*	
	4.6	5.1	3.6	*1.7*		2.8	2.3	2.4	*7.5*	
	4.3	5.0	3.8	2.0		3.8	1.5	1.4	4.9	
	3.6	4.6	4.0	2.8		7.0	3.3	2.8	3.4	
8	10.6	2.4	0.2	0.0		3.6	0.6	0.1	0.0	
	8.1	5.9	0.8	0.0		9.8	4.5	0.5	0.1	
	5.1	*6.6*	2.8	0.5		6.9	*8.7*	3.5	0.7	
	3.9	5.9	*3.9*	1.3		7.5	7.3	*5.9*	2.4	
9	10.6	2.7	0.3	0.0		2.2	0.6	0.2	0.0	
	8.9	4.7	0.8	0.1		*8.4*	4.2	0.6	0.2	
	5.3	*6.5*	2.7	0.5		7.6	*8.9*	3.9	0.9	
	3.8	5.4	*4.3*	1.5		7.5	6.6	*7.2*	2.9	
10	9.8	3.2	0.5	0.1		2.6	1.8	0.6	0.4	
	6.2	*6.2*	2.2	0.3		5.1	*6.5*	2.2	1.3	
	5.5	5.5	*3.1*	0.9		5.3	4.7	*4.9*	3.6	
	4.4	4.9	3.8	*1.9*		8.3	4.3	3.9	*5.4*	

Table 4. Total number of observed toxicities for the restricted CRM and Bayesian copula design under the first 10 scenarios

Scenario	*Results for the restricted CRM*	*Results for the Bayesian copula design*
1	26.2	20.3
2	23.5	21.8
3	17.3	20.4
4	20.5	21.3
5	8.4	14.5
6	27.8	8.2
7	14.8	17.5
8	24.0	22.2
9	23.7	22.3
10	20.9	20.7

1, the restricted CRM treated the majority of patients at dose combinations above the MTD, whereas in scenario 5 most of the patients were treated at doses below the MTD. In addition to the benefit of treating more patients at desirable doses, our method is safer than the restricted CRM in terms of the average number of patients experiencing toxicity in six of the 10 cases, as shown in Table 4. In scenario 5, on average, six more patients experienced toxicity under our design than under the restricted CRM, as most of the patients were treated at the under-MTD combinations in the restricted CRM design. Under scenario 6, in which all doses were overly toxic, the restricted CRM had to be implemented independently for each of the four rows as it

Table 5. Sensitivity analysis of the Bayesian copula design, including model misspecifications and alternative prior specifications

Results for selection probability					*Results for numbers of patients treated*				
Scenario 11 (Frank copula)									
1.9	0.5	0.0	0.0	0.0	1.1	0.3	0.1	0.0	0.0
8.7	3.0	0.7	0.2	0.1	5.1	1.4	0.3	0.1	0.0
24.1	12.6	3.7	0.5	0.3	*14.0*	5.2	1.2	0.2	0.1
2.5	*20.4*	6.9	1.1	0.4	13.6	*9.9*	3.7	1.1	0.3
Scenario 12 (Gumbel copula)									
5.6	1.7	0.2	0.0	0.0	2.3	0.9	0.1	0.0	0.0
14.2	9.5	1.6	0.2	0.0	*7.3*	3.1	0.7	0.1	0.0
9.3	*18.8*	8.5	1.4	0.1	10.9	*6.8*	2.8	0.6	0.1
0.7	6.7	*8.6*	3.5	0.3	9.8	6.1	*4.6*	1.9	0.6
Scenario 3 with prior $\alpha, \beta \sim$ gamma(1, 1)									
3.4	*11.2*	9.5	1.2	0.2	3.3	*4.1*	2.7	0.9	0.4
1.1	7.6	*12.7*	7.6	1.8	3.9	4.0	*4.2*	2.2	1.5
0.0	0.7	4.2	*12.7*	9.8	4.1	1.3	2.5	*4.0*	3.7
0.0	0.1	0.2	3.6	*12.2*	6.4	2.9	2.7	3.3	*4.9*
Scenario 3 with prior $\alpha, \beta \sim$ gamma(0.5, 0.5)									
3.6	*11.4*	10.7	2.0	0.1	3.3	*4.0*	3.2	1.1	0.4
1.0	7.5	*12.4*	7.6	1.9	3.6	3.9	*4.6*	2.4	1.3
0.0	0.8	6.2	*11.6*	8.5	4.2	1.6	3.2	*3.9*	3.2
0.0	0.0	0.5	2.9	*10.8*	6.3	2.9	2.7	2.9	*4.2*

could not borrow information across the rows. As a result, the restricted CRM allocated at least a few patients to the first column and more patients experienced toxicity than in our design, which stopped the trial earlier. As to other scenarios, the method proposed performed comparably with the restricted CRM. Across the 10 scenarios, on average 207.1 patients experienced toxicity by using the restricted CRM as opposed to 189.2 patients by using our method.

We conducted two types of sensitivity analysis, which are summarized in Table 5. To examine the robustness of our method to model misspecifications, the probabilities of toxicity in scenarios 11 and 12 were simulated from the Frank and Gumbel copula models (comparing with scenarios 1 and 2) respectively. Under these two scenarios, our method performed very well with high selection percentages of the MTD combinations. Moreover, to evaluate the effect of the prior specifications, we investigated the performance of our design under two more diffusive prior distributions for α and β under scenario 3. The Bayesian copula design appeared not to be sensitive to the prior specification and yielded very similar results on the basis of different hyperparameters.

5. Concluding remarks

To accommodate the enormous need for designing clinical trials with drug combinations, we have developed a copula-type model to link the joint toxicity probability with the probabilities of toxicity of each drug. The method proposed can fully evaluate the joint toxicity profiles of the combined drugs, as well as preserve their single-agent properties. The drug–drug interactive effects are naturally modelled through a copula-type model, which reduces to the CRM design if only one drug is considered. The attractive feature of this design is that it efficiently reorders the probabilities of toxicity of the dose combinations on the basis of the accrued data, so that the next cohort of patients to be treated will receive the most appropriate dose. The toxicity surface for the combined drugs can be adequately captured and reshaped by the three unknown parameters. In a typical drug combination trial, the doses of each drug are often bounded by the corresponding single-agent MTDs for which the toxicity probabilities are known from previous studies. Therefore, the prior specification for the probabilities of toxicity of each drug are much more accurate than the usual CRM for single-agent cases. The start-up rule was introduced for safety and did not affect the performance of the trial. We also implemented the design proposed by using the Gumbel copula in the simulation study, and we obtained similar results to those in the Clayton copula. The Bayesian Markov chain Monte Carlo algorithm can coherently update the posterior estimates for the model parameters as more patients enter the trial and more outcomes are observed. It achieves the goal of designing a Bayesian dose finding clinical trial by borrowing strength from all the available data.

Acknowledgements

We thank the referees, Associate Editor and Joint Editor for helpful comments that substantially improved the paper. The research was partially supported by funds from the Physician Referral Service at the M. D. Anderson Cancer Center, 5 P50 CA116199-03 breast cancer 'Specialized program of research excellence' and US Department of Defense grant W81XWH-05-2-0027.

References

Abdelbasit, K. M. and Plackett, R. L. (1982) Experimental design for joint action. *Biometrics*, **38**, 171–179.
Ashford, J. R. (1981) General models for the joint action of mixtures of drugs. *Biometrics*, **37**, 457–474.
Babb, J., Rogatko, A. and Zacks, S. (1998) Cancer phase I clinical trials: efficient dose escalation with overdose control. *Statist. Med.*, **17**, 1103–1120.

Braun, T. M., Thall, P. F., Nguyen, H. and de Lima, M. (2007) Simultaneously optimizing dose and schedule of a new cytotoxic agent. *Clin. Trials*, **4**, 113–124.
Chevret, S. (2006) *Statistical Methods for Dose-finding Experiments*. Chichester: Wiley.
Clayton, D. G. (1978) A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, **65**, 141–152.
Conaway, M. R., Dunbar, S. and Peddada, S. D. (2004) Designs for single- or multiple-agent phase I trials. *Biometrics*, **60**, 661–669.
Durham, S. D., Flournoy, N. and Rosenberger, W. F. (1997) A random walk rule for phase I clinical trials. *Biometrics*, **53**, 745–760.
Edler, L. (2001) Overview of phase I trials. In *Handbook of Statistics in Clinical Oncology* (ed. J. Crowley), pp. 1–34. New York: Dekker.
Gasparini, M. and Eisele, J. (2000) A curve-free method for phase I clinical trials. *Biometrics*, **56**, 609–615.
Genest, C. and Rivest, L.-P. (1993) Statistical inference procedures for bivariate Archimedean copulas. *J. Am. Statist. Ass.*, **88**, 1034–1043.
Goodman, S. N., Zahurak, M. L. and Piantadosi, S. (1995) Some practical improvements in the continual reassessment method for phase I studies. *Statist. Med.*, **14**, 1149–1161.
Hougaard, P. (1986) A class of multivariate failure time distributions. *Biometrika*, **73**, 671–678.
Huang, X., Biswas, S., Oki, Y., Issa, J. P. and Berry, D. A. (2007) A parallel phase I/II clinical trial design for combination therapies. *Biometrics*, **63**, 429–436.
Korn, E. L., Midthune, D., Chen, T. T., Rubinstein, L. V., Christian, M. C. and Simon, R. M. (1994) A comparison of two phase I trial designs. *Statist. Med.*, **13**, 1799–1806.
Korn, E. L. and Simon, R. (1992) Selecting dose-intense drug combinations: metastatic breast cancer. *Breast Cancer Res. Trtmnt*, **20**, 155–166.
Korn, E. L. and Simon, R. (1993) Using the tolerable-dose diagram in the design of Phase I combination chemotherapy trials. *J. Clin. Oncol.*, **11**, 794–801.
Kramar, A., Lebecq, A. and Candalh, E. (1999) Continual reassessment methods in phase I trials of the combination of two drugs in oncology. *Statist. Med.*, **18**, 1849–1864.
Kuzuya, K., Ishikawa, H., Nakanishi, T., Kikkawa, F., Nawa, A., Fujimura, H., Iwase, A., Arii, Y., Kawai, M., Hattori, S., Sakakibara, K., Sasayama, E., Furuhashi, Y., Suzuki, T. and Mizutani, S. (2001) Optimal doses of paclitaxel and carboplatin combination chemotherapy for ovarian cancer: a phase I modified continual reassessment method study. *Int. J. Clin. Oncol.*, **6**, 271–278.
Leung, D. H.-Y. and Wang, Y.-G. (2002) An extension of the continual reassessment method using decision theory. *Statist. Med.*, **21**, 51–63.
Lokich, J. (2001) Phase I clinical trial of weekly combined topotecan and irinotecan. *Am. J. Clin. Oncol.*, **24**, 336–340.
Møller, S. (1995) An extension of the continual reassessment methods using a preliminary up-and-down design in a dose finding study in cancer patients, in order to investigate a greater range of doses. *Statist. Med.*, **14**, 911–922.
Mukhopadhyay, S. (2000) Bayesian nonparametric inference on the dose level with specified response rate. *Biometrics*, **56**, 220–226.
Nelsen, R. B. (1999) *An Introduction to Copulas*. New York: Springer.
O'Quigley, J., Paoletti, X. and Maccario, J. (2002) Non-parametric optimal design in dose finding studies. *Biostatistics*, **3**, 51–56.
O'Quigley, J., Pepe, M. and Fisher, L. (1990) Continual reassessment method: a practical design for Phase 1 clinical trials in cancer. *Biometrics*, **46**, 33–48.
O'Quigley, J. and Shen, L. Z. (1996) Continual reassessment method: a likelihood approach. *Biometrics*, **52**, 673–684.
Plackett, R. L. and Hewlett, P. S. (1967) A comparison of two approaches to the construction of models for quantal responses to mixtures of drugs. *Biometrics*, **23**, 27–44.
Rosenberger, W. F. and Haines, L. M. (2002) Competing designs for phase I clinical trials: a review. *Statist. Med.*, **21**, 2757–2770.
Simon, R. and Korn, E. (1990) Selecting drug combinations based on total equivalent dose (dose intensity). *J. Natn. Cancer Inst.*, **82**, 1469–1476.
Storer, B. E. (1989) Design and analysis of Phase I clinical trials. *Biometrics*, **45**, 925–937.
Stylianou, M. and Flournoy, N. (2002) Dose finding using the biased coin up-and-down design and isotonic regression. *Biometrics*, **58**, 171–177.
Thall, P. F., Millikan, R. E., Müller, P. and Lee, S.-J. (2003) Dose-finding with two agents in phase I oncology trials. *Biometrics*, **59**, 487–496.
Wang, K. and Ivanova, A. (2005) Two-dimensional dose finding in discrete dose space. *Biometrics*, **61**, 217–222.
Whitehead, J. and Brunier, H. (1995) Bayesian decision procedures for dose determining experiments. *Statist. Med.*, **14**, 885–893.
Yin, G. and Yuan, Y. (2008) A latent contingency table approach to dose finding for combinations of two agents. *Biometrics*, to be published.