DOSE FINDING METHODS BASED ON CURE RATE MODEL IN PHASE I CANCER CLINICAL TRIALS

By

MENGHUI CHEN

A dissertation submitted to the

School of Public Health

and the

Graduate School — New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Written under the direction of

Professor Yong Lin

And Approved by

_____

_____

_____

_____

New Brunswick, New Jersey

May, 2015

ABSTRACT OF THE DISSERTATION

Dose Finding Methods Based on Cure Rate Model in Phase I Cancer Clinical Trials

by Menghui Chen

Dissertation Director: Professor Yong Lin

The main goal of a Phase I cancer clinical trial is to identify the maximum tolerated dose (MTD) of a new drug having acceptable dose-limiting toxicity (DLT). Two main model-based designs are continual reassessment method (CRM) (O'Quigley et al., 1990) and escalation with overdose control (EWOC) (Babb et al., 1998). Most of the designs are based on the binary toxic outcome. The occurrence of DLT is assessed over a predefined time window, and complete follow-up of the current patient is required to fit the model. Information is lost by categorizing time to DLT to a binary variable and might lead to a poor estimate of MTD. Trials might have to suspend accrual to obtain complete data and lead to long trial durations and complicate administrative burdens. Some methods have been proposed to incorporate the time-to-DLT using a weight function, such as TITE-CRM by Cheung and Chappell (2000) and TITE-EWOC by Mauguen et al. (2011). A better approach would be to model the time-to-DLT data directly for patients who will experience the DLT and to separate these patients from those who will never experience the DLT given a specific dose. This approach can be based on the well-studied cure model, a type of mixture model. This mixture model seems to be more appropriate for dose finding in Phase I cancer studies when time-to-DLT is incorporated. In this thesis, we will develop a Bayesian design framework based

on cure model approach to incorporate the time-to-DLT and will extend the current model-based designs such as CRM, EWOC or the hybrid design (Chu et al., 2009) to incorporate the time-to-DLT event. We will call such design as CATE design for Cure rate model Approach for Time-to-DLT Event. To evaluate performance of CATE designs, extensive simulation studies had been conducted, and the results shows that CATE designs outperform the existing designs for phase I cancer clinical trials.

Acknowledgements

I would like to express my sincere appreciation and thanks to my advisor Professor Dr. Yong Lin, who has been a tremendous mentor for me. His enthusiasm and immense knowledge guided me through all these years of research and writing of this thesis. I am grateful for his patience and encouragement which motivates me to strive towards my goal and have made this journey much more enjoyable.

I would especially like to thank the rest of my committee members, Professor Weichung Joe Shih, Professor Sinae Kim, and Dr. Yujun Wu for their insightful comments and suggestions, perfecting the ideas and writing of this thesis. I would also like to thank other faculty members in the Biostatistics department for their encouragement and support during my Ph.D. study.

Last but not least, I would like to thank my family for their love and support to make this sustained effort possible.

Dedication

To my parents

Qinghua Chen and Liling Wu

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

The main goal of a Phase I cancer clinical trial is to identify the maximum tolerated dose (MTD) of a new drug having acceptable dose-limiting toxicity (DLT). Typically the MTD is defined as the highest dose at which a predefined target percentage of patients experience a DLT. The highest dose is sought since the benefit of the new treatment is believed to increase with dose. The DLT should be defined prior to the enrollment of patients in the trial. In the United States, the NCI (National Cancer Institute) common toxicity criteria is usually used to define DLT as a group of toxicities of grade three or higher. The grades are defined as follows: grade 0, no toxicity; grade 1, mild; grade 2, moderate; grade 3, severe; grade 4, life-threatening; grade 5, death.

Phase I trials involve humans and must adhere to the ethical norms of clinical research. As first-in-human studies, the safety of the participants is of primary concern. The ethical issues have influenced the sample size and the design of these studies. The challenges in Phase I cancer trial design and conduct are that limited information is known about the relationship between dose and probability of toxicity before the trial begins, but decisions must be based on very small sample sizes. Unlike phase I trials in many medical areas where the expected toxicity is mild and suboptimal doses of new agents or combination therapies are administered to healthy volunteers, phase I cancer trials of new agents or combination therapies cannot be conducted in healthy volunteers due to the toxicity that generally observed in preclinical studies. The participants in phase I cancer clinical trials are almost always patients at advanced disease stages who have not responded to the standard treatment and consent to participate in the trial for seeking a cure as their last resort. These patients are often at very high risk of death in a relatively short term. For ethical reasons, phase I cancer clinical trials must balance

between increasing the dose gradually to ensure the safety of patients and minimizing the number of patients treated at low doses that may have no therapeutic effect. Therefore, randomly assigning patients between several dose levels appear unacceptable. Sequential entering patients to gradually increase doses by some dose-escalation scheme is often used. An efficient design should maximize the proportion of patients assigned to the optimal dose and use no more patients than necessary.

Various designs have been developed for dose-finding in cancer studies in the literature. One commonly used method is the algorithm-based design, such as the traditional "3+3" design (Edler, 1990; Lin and Shih, 2001; Shih and Lin, 2006). Three or six patients are treated at each dose level, depending on the observed toxicity until a specified number of DLT incidence is observed at a dose level. Despite the fact that the traditional method for dose escalation has been criticized for its tendency to include too many patients at suboptimal dose levels and give a poor estimate of the MTD (O'Quigley et al., 1990; Heyd and Carlin, 1990), it is still widely used in practice because of its algorithm-based simplicity in logistics for the clinical investigators to carry out. Lin and Shih (2001) discussed key statistical properties of the traditional and modified algorithm-based designs in a general framework of "A+B" designs and derived the exact formulas for the corresponding statistical properties. The statistical properties include: (1) the probability of a dose being chosen as the MTD; (2) the expected number of patients at each dose level; (3) the target toxicity level (the expected DLT at the MTD found through the algorithm-based design); (4) the expected DLT incidence at each dose level; (5) the expected overall DLT incidences in the trial. They showed that the algorithm-based design does not have a fixed target toxicity level (such as 33% for the "3+3" design in many people's misconception).

Another important class of the designs for Phase I cancer studies is the model-based designs, where some form of a monotonic dose-response curve is assumed and MTD can be estimated from all accumulated data. Two main model-based designs are continual reassessment method (CRM) (O'Quigley et al., 1990) and escalation with overdose control (EWOC) (Babb et al., 1998). CRM and EWOC are adaptive designs. The dose-response relationship is updated as observations on DLT become available.

In CRM, patients are always treated at the dose whose probability of a DLT, according to the current knowledge, is closest to the desired level. It was originally developed in Bayesian framework, although the method was also developed based on likelihood (CRML). It has been demonstrated that CRM outperforms the algorithm-based design such as "3+3" and up-and-down designs with increased efficiency and precision together with a lower bias. EWOC is designed to approach the MTD as fast as possible subject to the constraint that the predicted proportion of patients who receives an overdose does not exceed a specified value (feasibility bound). It has lower over-dose proportion of patients relative to CRM. But because of the overdose control in EWOC, the dose escalation to MTD may not be as quick as in CRM. Several modifications have been made to CRM and EWOC, and a hybrid design to unify CRM and EWOC has been proposed by Chu et al. (2009). The hybrid design is EWOC with gradually increasing feasibility bound to 0.5 (median), which becomes a modified CRM when reaching the feasibility bound of 0.5. The hybrid design has been shown to have faster convergence over EWOC and better safety protection over CRM.

Most of the designs are based on the binary toxic outcome. The occurrence of DLT is assessed over a predefined time window, and complete follow-up of the current patient is required before the next patient is assigned. There are several issues with this approach. First, information is lost by categorizing time to DLT to a binary variable. Second, if the follow-up time is too short, especially in the case of late-onset toxicity, the DLT probability might be underestimated and leads to a poor estimate of MTD. Third, the trial might have to suspend accrual until all the patients in the current cohort complete their follow-up. This may lead to long trial durations and complicate administrative burdens. Cheung and Chappell (2000) proposed a method to incorporate the time-to-DLT into the CRM (TITE-CRM). They considered a weighted dose-response model, where the weight is a function of time-to-DLT of a patient. The weight reflects the information available from a patient, and it equals to one if a DLT is observed or to the proportion of time until the full assessment time if the patient has no DLT and not fully observed. Their simulation showed that the performance of TITE-CRM are comparable with CRM's but TITE-CRM takes a much shorter trial duration. Similar

approach has been adapted to the EWOC method by Mauguen et al. (2011). They showed that TITE-EWOC greatly decreased the trial duration while maintaining the overdose control and the performance of the EWOC method. These approaches did not fully incorporate time-to-DLT into the model and patients with different time to DLT contribute the same to the likelihood. It assumes that the hazard of toxicity remains constant over time, which may not be the case. In addition, they did not consider that some patients never develop a DLT given a dose.

A better approach would be to model the time-to-DLT data directly for patients who will experience the DLT and to separate these patients from those who will never experience the DLT given a specific dose. This approach can be based on the well-studied cure model (Boag, 1949; Gamel et al., 1990; Yamaguchi, 1992; Sy and Taylor, 2000a; Peng, 2003; Wu et al., 2014), a type of mixture model. This mixture model seems to be more appropriate for dose finding in Phase I cancer studies when time-to-DLT is incorporated.

In this thesis, we will develop a Bayesian design framework based on cure model approach to incorporate the time-to-DLT event. We will extend the current model-based designs such as CRM, EWOC or Hybrid design to incorporate the time-to-DLT event. We will call such designs as CATE designs to stand for Cure rate model Approach for Time-to-DLT Event.

Chapter 2

Literature Review

2.1 Continual Reassessment Method (CRM)

The continual reassessment method (CRM) was first proposed by O'Quigley et al. (1990) as an alternative method of the algorithm-based designs in phase I cancer clinical trials. Its basic assumption is that the probability of toxic response increases monotonically with increasing dose. The authors argued that for these trials the particular shape of the dose toxicity curve is of little interest, and the only requirements for the models are that locally (i.e., around the dose corresponding to the targeted toxicity level) they reasonably well approximate the true probability of toxic response.

The CRM is an adaptive design and uses all of the available data prior to the trial and all the cumulative data from the trial to determine the dose level for new patients enrolled. The dose-response relationship is continually updating as observations on severe toxicity become available. Patients are always treated at the dose whose probability of toxicity, according to the current knowledge, is closest to the desired level.

Consider a set of prespecified dose levels to be tested: $\Omega_d = \{d_i : i = 1, ..., k, d_1 < d_2 < ... < d_k\}$. Let Y_j be the binary random variable of the DLT response, where $Y_j = 1$ denotes the DLT response, and $Y_j = 0$ denotes no DLT for the jth patient. Let x_j be the treatment dose level assigned to the jth patient, where x_j can be any dose levels $\{d_1, ..., d_k\}$. Let θ be the target probability of DLT. Reasonable target toxic probability can be chosen depending on types of drugs and the nature of DLT. We would set it relatively high when the DLT is a transient, correctable or non-fatal condition, and low when it is lethal or life threatening. The original CRM paper used $\theta = 0.20$. The objective of the phase I study is to estimate the maximum tolerated dose (MTD), where the probability of DLT at MTD is equal to the target toxicity level θ. The relationship

is described by the following equation:

$$P(Y = 1 \mid x = MTD) = \theta. \tag{2.1}$$

Consider some simple dose-response model $\psi(x_i, a)$:

$$P(Y = 1 \mid x_i) = \psi(x_i, a). \tag{2.2}$$

The only assumption is that $\psi(., a)$ is a monotonic increasing function in x_i and a and for any θ, there exists a unique a, say a_0 from $(0, \infty)$, such that $\psi(x_{MTD}, a_0) = \theta$. This parametric model should be flexible enough to approximate the underlying true dose-toxicity relationships in the neighborhood of the target toxic probability. O'Quigley et al. (1990) used the hyperbolic tangent function as their working model:

$$\psi(x_i, a) = \{(\tanh x_i + 1)/2\}^a. \tag{2.3}$$

Other working models have been proposed in the literature. For example, O'Quigley and Chevret (1991) suggested the one parameter logistic function:

$$\psi(x_i, \beta) = \frac{\exp(\alpha + \beta x_i)}{1 + \exp(\alpha + \beta x_i)}, \tag{2.4}$$

with the intercept α as a fixed constant.

Let $\mathcal{H}_j = \{(x_1, y_1), ..., (x_{j-1}, y_{j-1})\}$ be the history of dose assignments toxicity responses for the first $j - 1$ patients, and let $f(\beta, \mathcal{H}_j)$ be a nonnegative function summarizing all available information about the parameter β such that $\int_0^\infty f(\beta, \mathcal{H}_j)d\beta = 1$, $j = 1, ..., n$. This is the current prior before the experimentation on the jth patient. To determine the dose level for the jth patient, we will need the estimates of the probability of DLT at each discrete dose level i. Let θ_{ij} be the mean DLT probability at dose level i for the jth patient given the accumulated information on the first $j - 1$ patients,

where

$$\theta_{ij} = \int_0^\infty \psi(d_i, \beta) f(\beta, \mathcal{H}_j) d\beta, \quad i = 1, ..., k. \tag{2.5}$$

Instead of working with the expected values of the probabilities over the space of parameter β, the authors also proposed an alternative estimate based on the expected values of β as described below:

$$\theta'_{ij} = \psi(d_i, \beta(j)), \quad i = 1, ..., k, \text{ where } \beta(j) = \int_0^\infty a f(a, \mathcal{H}_j) da. \tag{2.6}$$

The authors used $\Delta(v, w)$ as some measure of distance between v and w, for example, $\Delta(v, w) = (v - w)^2$. For the jth patient, we can assign the dose level d_i that minimizes $\Delta(\theta_{ij}, \theta)$, $\Delta(\theta'_{ij}, \theta)$, or $\Delta(d_i, \psi^{-1}_{a=a(j)}(\theta))$, depending on the criterion we choose to work with. The first criterion requires performing k infinite integrals while the latter two only need one single integral. The latter two criteria are more computationally economic.

After determining the treatment level for the jth patient to x_j, and observing the response y_j, we can then update our knowledge about the parameter a_0. By Bayes theorem, the posterior density $f(a, \mathcal{H}_{j+1})$ can be obtained as:

$$f(a, \mathcal{H}_{j+1}) = \frac{\phi(x_j, y_j, a) f(a, \mathcal{H}_j)}{\int_0^\infty \phi(x_j, y_j, u) f(u, \mathcal{H}_j) du} = \frac{g(a) \prod_{l=1}^j \phi(x_j, y_j, a)}{\int_0^\infty g(u) \prod_{l=1}^j \phi(x_j, y_j, a) du}, \tag{2.7}$$

where $\phi(x_j, y_j, a) = \psi^{y_j}(x_j, a)\{1 - \psi(x_j, a)\}^{1-y_j}$, and $g(a) = f(a, \mathcal{H}_1)$. $g(a)$ is the initial prior distribution for β before the trial begins. The original CRM paper used a standard exponential prior because of its simplicity and its positivity constraint.

Based on the posterior density $f(a, \mathcal{H}_{j+1})$, CRM finds the next recommended dose level x_{j+1} for the $(j+1)$th patient by minimizing $\Delta(\theta_{ij+1}, \theta)$, $\Delta(\theta'_{ij+1}, \theta)$, or $\Delta(d_i, \psi^{-1}_{a=a(j+1)}(\theta))$. Continue in this way until the last patient is entered. The MTD can then be estimated after the response of the last patient is observed.

It has been demonstrated that the original CRM outperforms the algorithm-based designs with increased efficiency and precision with a lower bias. However, it has some difficulties to be implemented in real practice. Some modifications to the original CRM

have been proposed to address these concerns. For example, the lowest dose level in the trial was proposed to be the starting dose as in the traditional design (Faries, 1994; Korn et al., 1994), instead of the estimated MTD based on the prior knowledge as in the original design. Some stopping rule can also be imposed to end the trial early. For example, Korn et al. (1994) proposed to stop the CRM procedure after six patients are assigned to the same dose. Combined methods were proposed, splitting design into two stages: a first stage with an up-and-down design until the first toxicity is observed and then a second stage with the CRM using all information obtained at that point (Moller, 1995).

2.2 Escalation with Overdose Control Design (EWOC)

EWOC (Babb et al., 1998) is designed to approach the MTD as fast as possible while restricting the predicted proportion of patients exposed to overdose not exceeding a specified value.

Let X_{min} and X_{max} denote the minimum and maximum dose levels available to be tested in the trial. The dose levels are chosen in the prior belief that X_{min} is safe to patients and MTD is between X_{min} and X_{max}, i.e., $X_{min} \leq MTD \leq X_{max}$. Let Y_j be the binary random variable of the DLT response, where $Y_j = 1$ denotes the DLT response, and $Y_j = 0$ denotes no DLT for the jth patient. First patient is assigned dose X_{min}, and subsequent patients will be assigned dose between X_{min} and X_{max}. The dose level selected for the jth patients, x_j, $j = 1, .., n$, is

$$x_1 = X_{min}, x_j \in [X_{min}, X_{max}], \text{ for all } j = 1, ..., n. \tag{2.8}$$

The dose toxicity relationship is modeled as:

$$P(Y = 1 \mid dose = x) = F(\alpha + \beta x), \tag{2.9}$$

where F is a specified distribution function, and α and β are unknown parameters. It is assumed that $\beta > 0$ so that the probability of DLT is a monotonic increasing function

of dose. The MTD is the dose level, x_θ, at which the probability of DLT is the target level θ :

$$P(Y = 1 \mid dose = x_\theta) = F(\alpha + \beta x_\theta) = \theta. \tag{2.10}$$

Let $\mathcal{H}_j = \{(x_1, y_1), ..., (x_j, y_j)\}$ be the history of dose assignments and toxicity responses for the first j patients. The likelihood function of (α, β) given $\mathcal{H}_j$ is

$$L(\alpha, \beta \mid \mathcal{H}_j) = \prod_{i=1}^{j} F(\alpha + \beta x_i)^{y_i}[1 - F(\alpha + \beta x_i)]^{1-y_i}. \tag{2.11}$$

The prior information about (α, β) is incorporated through a prior distribution $h(\alpha, \beta)$ defined on

$$\Phi = \{(a, b) \in \mathcal{R}^2 : b > 0, F(a + bX_{min}) \leq \theta \leq F(a + bX_{max})\}. \tag{2.12}$$

By Bayes theorem the joint posterior distribution of (α, β) given data $\mathcal{H}_j$ is

$$P(\alpha, \beta \mid \mathcal{H}_j) = \tau^{-1} L(\alpha, \beta \mid \mathcal{H}_j) h(\alpha, \beta) I_\Phi(\alpha, \beta), \tag{2.13}$$

where

$$\tau = \iint_\Phi L(x, \mid \mathcal{H}_j) h(x, y) dx dy,$$

and $I_\Phi(\alpha, \beta)$ is the indicator function for the set Φ.

Babb et al. (1998) have re-parameterized the model in terms of MTD (x_θ) and the probability of DLT at the starting dose (ρ_0), where

$$\rho_0 = P(Y = 1 \mid X_{min}),$$

$$x_\theta = \frac{F^{-1}(\theta) - \alpha}{\beta} = X_{min} + \frac{F^{-1}(\theta) - F^{-1}(\rho_0)}{\beta}.$$

These parameters are easier to interpret to clinicians. The MTD (x_θ) is the parameter of interest, and one often have more information about the starting dose through preliminary studies so that a meaningful informative prior for ρ_0 can be specified. The

marginal posterior cumulative distribution function (CDF) of MTD given $\mathcal{H}_j$ can be derived from (2.13) through the transformation $T(\alpha, \beta) = (\rho_0, x_\theta)$. Denote the image of Φ under the transformation of T by $T(\Phi)$ and it follows that

$$T(\Phi) = [0, \theta] \times [X_{min}, X_{max}]. \tag{2.14}$$

The inverse transformation is

$$T^{-1}(\rho_0, x_\theta) = (f_1(\rho_0, x_\theta), f_2(\rho_0, x_\theta)), \tag{2.15}$$

where the functions f_1 and f_2 are defined on $T(\Phi)$ by

$$f_1(\rho_0, x_\theta) = \frac{x_\theta F^{-1}(\rho_0) - X_{min} F^{-1}(\theta)}{x_\theta - X_{min}}, \tag{2.16}$$

$$f_2(\rho_0, x_\theta) = \frac{F^{-1}(\theta) - F^{-1}(\rho_0)}{x_\theta - X_{min}}. \tag{2.17}$$

The joint posterior distribution of (ρ_0, x_θ) given $\mathcal{H}_j$ can be written as:

$$P(\rho_0, x_\theta \mid \mathcal{H}_j) = \tau^{-1} L(f_1(\rho_0, x_\theta), f_2(\rho_0, x_\theta) \mid D_j) g(\rho_0, x_\theta), \tag{2.18}$$

here $g(\rho_0, x_\theta)$ is the prior distribution induced for (ρ_0, x_θ) by the choice of $h(\alpha, \beta)$. One can choose to specify the prior $g(\rho_0, x_\theta)$ directly, rather than indirectly through $h(\alpha, \beta)$.

Let $\Theta(x_\theta) = \{\rho_0 : (\rho_0, x_\theta) \in T(\Phi)\}$. The marginal posterior PDF and CDF of the MTD given $\mathcal{H}_j$ can be written as

$$\pi(x_\theta | \mathcal{H}_j) = \iint_{\Theta(x_\theta)} P(\rho_0, x_\theta \mid \mathcal{H}_j) d\rho_0, \tag{2.19}$$

and

$$\pi_j(z) = \int_{X_{min}}^{z} \pi(x_\theta | \mathcal{H}_j) dx_\theta, \ z \in [X_{min}, X_{max}]. \tag{2.20}$$

EWOC can be described as follows. The first patient receives the dose $x_1 = X_{min}$. Each subsequent patient receives the dose so that based on all available data the posterior probability of exceeding the MTD is equal to the feasibility bound γ. For the jth patient, EWOC selects the dose such that

$$\pi_{m(j)}(x_j) = P(MTD \leq x_j \mid \mathcal{H}_{m(j)}) = \gamma, \tag{2.21}$$

where $m(j)$ is the number of observations available at the time of treatment for the jth patient. Hence, the dose level for the jth patient is

$$x_j = \pi^{-1}_{m(j)}(\gamma). \tag{2.22}$$

The dose sequence calculated from equation (2.22) assumes that all dose levels between X_{min} and X_{max} are available for the trial. In practice, phase I clinical trials are typically based on a set of prespecified dose levels $\Omega_d = \{d_i : i = 1, ..., k, d_1 < ... < d_k\}$. In this case, the jth patient receives the dose level that is closest to $\pi^{-1}_{m(j)}(\gamma)$. EWOC does not require that all patient responses are observed before treating a new patient. Instead, one can select the dose for the new patient based on the current available data.

Upon completion of the trial, the MTD is estimated by minimizing the posterior expected loss with respect to some loss function l. Asymmetric loss function is recommended since underestimation and overestimation have very different consequences. The dose selection by EWOC actually corresponds to the following asymmetric loss function:

$$l_\alpha(x, x_\theta) = \begin{cases} \gamma(x_\theta - x) & \text{if } x \leq x_\theta, \text{ i.e., if } x \text{ is an underdose} \\ (1-\gamma)(x - x_\theta) & \text{if } x > x_\theta, \text{ i.e., if } x \text{ is an overdose.} \end{cases} \tag{2.23}$$

In their simulation study, Babb et al. (1998) used the logistic to model dose toxicity relationship and uniform prior distribution for (ρ_0, x_θ). The feasibility bound γ is set to 0.25, and the target probability of DLT θ is 1/3. The results indicated that EWOC decreased the over-dose proportion of patients and exhibited fewer DLT relative to

CRM.

2.3 A Hybrid Design of CRM and EWOC

The advantage of CRM is faster convergence, while the advantage of EWOC is safety protection. Chu et al. (2009) showed that CRM and EWOC designs can be unified and proposed a hybrid design to have faster convergence over EWOC and better safety protection over CRM. The authors argued that the overdose control is more desirable at the beginning of the trial than at the end of the trial. The approach of the hybrid design is to gradually increase the feasibility bound of EWOC to 0.5 (median), which becomes a modified CRM when reaching the highest feasibility bound.

Consider a set of prespecified dose levels to be tested: $\Omega_d = \{d_i : i = 1, ..., k, d_1 < ... < d_k\}$. Let Y_j be the binary random variable of the DLT response and x_j be the treatment dose level assigned to the jth patient, where $x_j \in \{d_1, ..., d_k\}$. Let $\mathcal{H}_j = \{(x_1, y_1), ..., (x_j, y_j)\}$ be the history of dose assignments toxicity responses for the first j patients, $j = 1, ..., n$.

The authors used a two-parameter dose-toxicity model:

$$P(Y_j = 1|x_j, \boldsymbol{\phi}) = F(\alpha + \beta x_j), \quad j = 1, ..., n, \tag{2.24}$$

where F is a specified distribution function, and α and β are unknown parameters. It is assumed that $\beta > 0$ so that the probability of DLT is a monotonic increasing function of dose. Define the parameter space of $\boldsymbol{\phi} = (\alpha, \beta)$ as

$$\phi = \{(\alpha, \beta) : [\alpha^*, \alpha^{**}] \times [\beta^*, \beta^{**}]\},$$

where $-\infty < \alpha^* < \alpha^{**} < \infty$ and $-\infty < \beta^* < \beta^{**} < \infty$. The logistic model and the probit model are two typical dose-response models:

$$F(\alpha + \beta x_j) = \frac{\exp(\alpha + \beta x_j)}{1 + \exp(\alpha + \beta x_j)}, \tag{2.25}$$

$$F(\alpha + \beta x_j) = \Phi(\alpha + \beta x_j), \tag{2.26}$$

where Φ is the cumulative distribution function of the standard normal distribution. Let θ be the target toxicity level, $0 < \theta < 1$. The MTD x_θ can be derived as

$$x_\theta = (F^{-1}(\theta) - \alpha)/\beta. \tag{2.27}$$

If F is the logistic model, then equation (2.27) becomes

$$x_\theta = (ln(\frac{\theta}{1-\theta}) - \alpha)/\beta. \tag{2.28}$$

Let $f(\boldsymbol{\phi}|\mathcal{H}_{j-1})$ be a nonnegative function summarizing all available information about the parameter $\boldsymbol{\phi}$ based on $\mathcal{H}_{j-1}$. This is the current prior before the experimentation on the jth patient. After observing the DLT response y_j of the jth patient treated at dose x_j, the posterior density of $\boldsymbol{\phi}$ can be derived as

$$f(\boldsymbol{\phi}, \mathcal{H}_j) = \frac{\phi(x_j, y_j, \boldsymbol{\beta}) f(\boldsymbol{\beta}, \mathcal{H}_{j-1})}{\int_{\mathcal{B}} \phi(x_j, y_j, \boldsymbol{u}) f(\boldsymbol{u}, \mathcal{H}_{j-1}) d\boldsymbol{u}} = \frac{g(\boldsymbol{\beta}) \prod_{l=1}^{j} \phi(x_j, y_j, \boldsymbol{\beta})}{\int_{\mathcal{B}} g(\boldsymbol{u}) \prod_{l=1}^{j} \phi(x_j, y_j, \boldsymbol{u}) d\boldsymbol{u}}, \tag{2.29}$$

where $\phi(x_j, y_j, \boldsymbol{\beta}) = F(\alpha + \beta x_j)^{y_j} \{1 - F(\alpha + \beta x_j)\}^{1-y_j}$, and $g(\boldsymbol{\beta}) = f(\boldsymbol{\beta}|\mathcal{H}_0)$, $\mathcal{H}_0 = \emptyset$, and $g(\boldsymbol{\beta})$ is the initial prior distribution for $\boldsymbol{\beta}$ before the trial begins.

The authors discussed how several CRM strategies and EWOC choose the dose level for the next patient and described the relationship between CRM and EWOC as follows.

CRM1: O'Quigley et al. (1990) proposed to choose the next recommended dose level so that the expected posterior probability of DLT over the parameter space is equal to the target level θ. Given the accumulated information on the first $j-1$ patients, the dose level for the jth patient is chosen at level x_j so that

$$E[F(\alpha + \beta x_j)|\mathcal{H}_{j-1}] = \int_{\mathcal{B}} F(\alpha + \beta x_j) f(\boldsymbol{\beta}, \mathcal{H}_{j-1}) d\boldsymbol{\beta} = \theta. \tag{2.30}$$

CRM2: O'Quigley et al. (1990) also proposed using the posterior mean $\hat{\boldsymbol{\beta}} = (\hat{\alpha}, \hat{\beta})$

to choose the next dose level, where $\hat{\boldsymbol{\beta}}$ is estimated based on the posterior density $f(\boldsymbol{\beta}, \mathcal{H}_{j-1})$. Given the accumulated information on the first $j-1$ patients, the dose level for the jth patient is

$$x_j = (F^{-1}(\theta) - \hat{\alpha})/\hat{\beta}. \tag{2.31}$$

CRM3: Chu et al. (2009) proposed this design as one of the modification of CRM, in order to make comparisons between EWOC and the hybrid designs. Based on the equation (2.27), the next recommended dose level x_j for the jth patient is chosen to be the expected value of $x_\theta = (F^{-1}(\theta) - \alpha)/\beta$, where the expectation is taken based on the posterior density $f(\boldsymbol{\beta}, \mathcal{H}_{j-1})$,

$$x_j = E[x_\theta|\mathcal{H}_{j-1}] = \int_{\mathcal{B}} \frac{F^{-1}(\theta) - \alpha}{\beta} f(\boldsymbol{\beta}, \mathcal{H}_{j-1}) d\boldsymbol{\beta}. \tag{2.32}$$

CRM4: Instead of using the posterior mean to determine the next dose level as in CRM1, we can use the posterior median as proposed by Shih et al. (1999). Given the accumulated information on the first $j-1$ patients, choose $x_j^{0.5}$ for the jth patient so that

$$P(F(\alpha + \beta x_j^{0.5}) \geq \theta|\mathcal{H}_{j-1}) = 0.5. \tag{2.33}$$

Equation (2.33) can be generalized to

$$P(F(\alpha + \beta x_j^{\gamma}) \geq \theta|\mathcal{H}_{j-1}) = \gamma, \tag{2.34}$$

where $\gamma \in (0, 1)$, and x_j^{γ} is the dose such that the posterior upper γ quantile of $F(\alpha + \beta x_j^{\gamma})$ is θ.

EWOC: Babb et al. (1998) proposed to choose the next dose level so that the posterior probability of exceeding the MTD x_θ is equal to the prespecified feasibility bound γ. Given the accumulated information on the first $j-1$ patients, choose x_j^{γ} for

the jth patient so that

$$P(x_j^\gamma \geq x_\theta | \mathcal{H}_{j-1}) = \gamma, \tag{2.35}$$

where the probability is calculated based on the posterior density $f(\boldsymbol{\beta}, \mathcal{H}_{j-1})$. Babb et al. (1998) used $\gamma = 0.25$.

Relationship between EWOC and CRM: Since $F(x)$ is an increasing function, we have

$$P(F(\alpha + \beta x_j^\gamma) \geq \theta | \mathcal{H}_{j-1}) = P(x_j^\gamma \geq x_\theta | \mathcal{H}_{j-1}) = \gamma. \tag{2.36}$$

When $\gamma = 0.5$, x_j^γ is the dose such that the posterior median of $F(\alpha + \beta x_j^\gamma)$ is θ, and x_j^γ is the same as $x_j^{0.5}$ in equation (2.33). Thus EWOC with $\gamma = 0.5$ is equivalent to CRM4, and CRM4 can be viewed as a reformulated form from EWOC.

The approach of the hybrid design is to increase the feasibility bound γ gradually from 0.1 to 0.5 for the first, say 1 to 30 patients. The smaller γ is used at the beginning of the trial when the safety protection is most needed due to the lack of dose-toxicity information. After entering 30 patients and enough information for the dose-toxicity model is obtained, the γ level is stayed at 0.5 to provide good convergence to the MTD. The increasing change of γ corresponds to the quantile change in $F(\alpha + \beta x_j^\gamma)$. When γ reaches 0.5, the hybrid design becomes CRM4 and the posterior median is used to determine the next dose level. Simulation results showed that the hybrid design generally has faster convergence rates than EWOC and smaller overdose proportions than CRM.

2.4 TITE-CRM and TITE-EWOC

Most of the phase I designs are based on the binary toxic outcome. The occurrence of DLT is assessed over a predefined time window, and complete follow-up of the current patient is required before the next patient is assigned. There are several issues with this approach. First, information is lost by categorizing time to DLT to a binary variable.

Second, if the follow-up time is too short, especially in the case of late-onset toxicity, the DLT probability might be underestimated and leads to a poor estimate of MTD. Third, the trial might have to suspend accrual until all the patients in the current cohort complete their follow-up. This may lead to long trial durations and complicate administrative burdens.

Cheung and Chappell (2000) proposed the time-to-event continual reassessment method (TITE-CRM) to incorporate time-to-DLT into the CRM and allow patients to be entered in a staggered fashion. They considered a weighted dose-response model, where the weight is a function of the actual assessment time of a patient. The weight reflects the information available from a patient, and it equals to one if a DLT is observed or the patient complete the full assessment time without a DLT.

Consider a set of prespecified dose levels to be tested: $\Omega_d = \{d_i : i = 1, ..., k, d_1 < ... < d_k\}$. Let Y_j be the binary random variable of the DLT response and x_j be the treatment dose level assigned to the jth patient, where $x_j \in \{d_1, ..., d_k\}$. The CRM assumes a parametric model $F(x, \beta)$ to describe the dose-toxicity relationship, where F is a monotonic increasing function in x. Cheung and Chappell (2000) considered a weighted dose-response model $G(x, w, \beta)$ that is monotone increasing in w with marginal constraints $G(x, 0, \beta) = 0$ and $G(x, 1, \beta) = F(x, \beta)$ for all x, β. The weight w is a function of the time-to-event of a patient. For simplicity, the authors incorporated the weight linearly into F: $G(x, w, \beta) = wF(x, \beta)$, for $0 \leq w \leq 1$. The parameter β can be estimated based on the weighted likelihood function:

$$\tilde{L}_n(\beta) = \prod_{j=1}^{n} G(x_j, w_{j,n}, \beta)^{y_{j,n}} [1 - G(x_j, w_{j,n}, \beta)]^{1-y_{j,n}}, \tag{2.37}$$

where $y_{j,n}$ and $w_{j,n}$ are the indication of DLT for the jth patient and the weight assigned to this observation just prior to the entry time of the $(n+1)$th patient, respectively.

Let τ be the planned assessment time window and u_j be the actual follow-up time of patient j when a new patient would enter the trial. The authors defined the weight

function as follows.

$$\begin{cases} w_j = \frac{u_j}{\tau} & \text{if } Y_j = 0 \\ w_j = 1 & \text{if } Y_j = 1. \end{cases} \tag{2.38}$$

The weight function should reflect the amount of information available from a patient. If a DLT is observed, the observation is complete and the weight should be 1. Other weight functions may also be considered as discussed in Cheung and Chappell (2000).

In their simulation study, Cheung and Chappell used the power function $F(x, \beta) = x^{\beta}$ as the dose-toxicity model and the standard exponential distribution function as the prior for β. Patients' failure time were generated under three models: a conditionally uniform model, a log-logistic model, and a Weibull model. The results showed that the performance of TITE-CRM are comparable with CRM's but TITE-CRM takes a much shorter trial duration.

Similar approach has been adapted to the EWOC method by Mauguen et al. (2011). They used the same weight function as in (2.38). In their simulation study, the authors used the logistic function for the dose-toxicity model

$$F(\alpha + \beta x_j) = \frac{\exp(\alpha + \beta x_j)}{1 + \exp(\alpha + \beta x_j)}. \tag{2.39}$$

The model is then re-parameterized in terms of MTD (x_θ) and the probability of DLT at the starting dose level (ρ_0). The priors for (x_θ, ρ_0) were chosen uniform and independent. Patients' failure time were generated using a uniform distribution only if a patient presented a DLT. The results showed that TITE-EWOC greatly decreased the trial duration while maintaining the overdose control and the performance of the EWOC method.

2.5 Cure Rate Model

Cure rate models are survival models that incorporates a cured fraction to model the non-zero tail probability of the survival function. Cure rate models have been used to model the time-to-event data when there is a significant proportion of patients expected

to be "cured", that is to remain disease-free even after long time follow-ups. One popular type of cure rate model is the mixture model discussed by Berkson and Gage (1952). This model separates the entire population into cured and uncured subjects. The survival function for the entire population can be written as:

$$S(t) = \pi S_L(t) + (1 - \pi), \tag{2.40}$$

where $S(t)$ denotes the survival function for the entire population, and $1 - \pi$ is the "cured fraction", and $S_L(t)$ is the survival function for the uncured subjects. The model in (2.40) is referred to as the standard cure rate model and has been studied by many authors.

Denote T as a non-negative random variable for the failure time, $\mathbf{x}$ and $\mathbf{z}$ as the covariate vectors, $\pi(\mathbf{z})$ as the uncured probability for a subject, and $S(t|\mathbf{x}, \mathbf{z})$ as the survival function for T, respectively. Let $f_L(t|\mathbf{x})$ and $S_L(t|\mathbf{x})$ be the probability density function (pdf) and the survival function for uncured subjects. In the presence of covariates, the cure model can be written in terms of the survival function:

$$S(t|\mathbf{x}, \mathbf{z}) = \pi(\mathbf{z}) S_L(t|\mathbf{x}) + [1 - \pi(\mathbf{z})]\,. \tag{2.41}$$

It is noted that $f(t|\mathbf{x}, \mathbf{z}) = -dS(t|\mathbf{x}, \mathbf{z})/dt = \pi(\mathbf{z}) f_L(t|\mathbf{x})$. Here the "incidence" part $\pi(\mathbf{z})$ is commonly modeled by logistic regression. The "latency" part $f_L(t|\mathbf{x})$ or $S_L(t|\mathbf{x})$ could be modeled parametrically, semi-parametrically, or non-parametrically. For parametric cure rate models, different distributions have been used to model the survival time for the uncured group, such as exponential, weibull, and generalized F distribution. For semi-parametric mixture models, Kuk and Chen (1992); Sy and Taylor (2000b); Peng et al. (1998) used semi-parametric proportional hazards models. An alternative semi-parametric approach uses accelerated failure time (AFT) models (Li and Taylor, 2002).

Chapter 3

Research Method

3.1 Design Based on Cure Model Approach for Time-to-DLT Event (CATE)

Denote T as a non-negative random variable for time-to-DLT and x as a dose level. The patient population is usually heterogeneous, it is a mixture of patients who will develop DLT (susceptible patients) and who will never develop DLT (unsusceptible patients). Given dose x, let $\pi(x)$ be the probability of a patient who will develop a DLT, and $1 - \pi(x)$ be the probability of a patient who will not develop a DLT. Let $S(t \mid x)$ and $S_L(t \mid x)$ be the DLT-free survival function of T for the overall population and susceptible patients, respectively, and $f_L(t \mid x)$ be the probability density function of T for susceptible patients. The cure model can be written as

$$S(t \mid x) = \pi(x)S_L(t \mid x) + (1 - \pi(x)). \tag{3.1}$$

$\pi(x)$ is called the "incidence" part and can be modeled by a logistic regression, i.e.,

$$\pi(x) \quad = \quad \frac{\exp(\alpha + \beta x)}{1 + \exp(\alpha + \beta x)}. \tag{3.2}$$

We assume that the proportion of susceptible patients $\pi(x)$ is an increasing function of dose, hence $\beta > 0$.

$S_L(t \mid x)$ is called the "latency" part and can be modeled by a proportional hazards model or an accelerated failure time model.

3.1.1 CATE Design Based on the Proportional Hazards Model

Under the proportional hazards model, the hazard function of DLT for a susceptible patient $h_L(t \mid x)$ can be written as

$$h_L(t \mid x) \quad = \quad h_0(t) \exp(\gamma x), \tag{3.3}$$

where $h_0(t)$ is the baseline hazard function corresponding to the risk of DLT for a susceptible patient given the standardized dose $x = 0$. The parameter γ represents the dose effect on the risk of DLT for patients who will develop a DLT. We assume that the hazard of DLT is an increasing function of dose, hence $\gamma > 0$.

Let $\mathcal{H}_n = \{(Y_i, \delta_i, x_i), i = 1, 2, ..., n\}$ be the observed data, where $Y_i = min(t_i, \tau_i)$ is the observed DLT time for the i^{th} patient who is observed up to time τ_i, δ_i is the censoring indicator with 1 if Y_i is the observed time-to-DLT, and 0 if censored. x_i is the standardized dose allocated to patient i. The likelihood is

$$L(\alpha, \beta, \gamma \mid \mathcal{H}_n) = \prod_{i=1}^{n} [\pi(x_i) f_L(t_i \mid x_i)]^{\delta_i} \{\pi(x_i) S_L(t_i \mid x_i) + [1 - \pi(x_i)]\}^{1-\delta_i}. \tag{3.4}$$

Assuming that the DLT-free survival time for the susceptible patients follow an exponential distribution with the baseline hazard function $h_0(t) = \lambda$ ($\lambda > 0$), then $h_L(t \mid x) = \lambda e^{\gamma x}$,

$$S_L(t \mid x) = \exp(-\lambda t e^{\gamma x}), \tag{3.5}$$

and the pdf of DLT for a susceptible patient is $f_L(t \mid x) = \lambda e^{\gamma x} \exp(-\lambda t e^{\gamma x})$. The likelihood (3.4) becomes

$$\begin{aligned} & L(\alpha, \beta, \gamma, \lambda \mid \mathcal{H}_n) \\ & = \prod_{i=1}^{n} [\pi(x_i) \lambda e^{-\gamma x_i} \exp(-\lambda t_i e^{\gamma x_i}]^{\delta_i} \{\pi(x_i) \exp(-\lambda t_i e^{\gamma x_i}) + [1 - \pi(x_i)]\}^{1-\delta_i}. \end{aligned}$$

We can also assume that the DLT-free survival time for the susceptible patients follows other parametric distributions. The commonly used distributions are exponential

(Jones et al., 1981; Goldman, 1984; Ghitany and Maller, 1992); Weibull (Farewell, 1982); Log-normal (Boag, 1949; Gamel et al., 1990); Gompertz (Gordon, 1990a,b); Extended generalized gamma (EGG) (Yamaguchi, 1992); and Generalized F (GF) distributions (Peng et al., 1998). We will consider both the exponential and Weibull distributions in this thesis.

If the survival time for the susceptible patients follow a Weibull distribution with the baseline hazard function $h_0(t) = v\lambda t^{(v-1)}$ ($\lambda > 0$, $v > 0$, λ is the rate parameter and v is the shape parameter), one can show that

$$h_L(t \mid x) = v\lambda t^{(v-1)} e^{\gamma x},$$

$$S_L(t \mid x) = \exp(-\lambda t^v e^{\gamma x}),$$

$$f_L(t \mid x) = v\lambda t^{(v-1)} e^{\gamma x} \exp(-\lambda t^v e^{\gamma x}).$$

The likelihood (3.4) becomes

$$\begin{aligned} & L(\alpha, \beta, \gamma, \lambda, v \mid \mathcal{H}_n) \\ &= \prod_{i=1}^{n} [\pi(x_i) v\lambda t_i^{(v-1)} e^{-\gamma x_i} \exp(-\lambda t_i^v e^{\gamma x_i})]^{\delta_i} \{\pi(x_i) \exp(-\lambda t_i^v e^{\gamma x_i}) + [1 - \pi(x_i)]\}^{1-\delta_i}. \end{aligned}$$

Priors and Posteriors

Assume that the DLT-free survival time for the susceptible patients follow an exponential distribution. Let $g(\alpha, \beta, \gamma, \lambda)$ be the prior distribution for α, β, γ, and λ. By Bayes' theorem the posterior given $\mathcal{H}_n$ is

$$G(\alpha, \beta, \gamma, \lambda \mid \mathcal{H}_n) \propto L(\alpha, \beta, \gamma, \lambda \mid \mathcal{H}_n) \times g(\alpha, \beta, \gamma, \lambda).$$

This posterior distribution will be used to determine the next dose for the next patient based on the original various designs, say CRM, EWOC or Hybrid (see O'Quigley et al. (1990); Babb et al. (1998); Chu et al. (2009)). In CRM, the next recommended

dose level is chosen so that the expected posterior probability of DLT over the parameter space is equal to the target level θ. In EWOC, the next dose level is chosen so that the posterior probability of exceeding the MTD x_θ is equal to the prespecified feasibility bound (e.g. 0.25). The approach of the hybrid design is to increase the feasibility bound gradually from 0.1 to 0.5 for the first, say 1 to 30 patients. When the feasibility bound reaches 0.5, the posterior median is used to determine the next dose level.

We assume that the prior distributions of $-\alpha$, β, γ and λ are independent such that $g(\alpha,\beta,\gamma,\lambda) = g_0(\alpha,\beta,\gamma)g_4(\lambda) = g_1(\alpha)g_2(\beta)g_3(\gamma)g_4(\lambda)$. The following prior distributions $g_0(\alpha,\beta,\gamma)$ for $-\alpha$, β and γ are considered:

- Exponential prior distribution

$$g_0(\alpha,\beta,\gamma) \;=\; (abc)\exp[-(a\alpha+b\beta+c\gamma)],$$

  where the means of the priors are $1/a$, $1/b$ and $1/c$ for α, β and γ, respectively.

- Gamma prior distribution

$$g_0(\alpha,\beta,\gamma) \;=\; \frac{a_2^{a_1}b_2^{b_1}c_2^{c_1}\alpha^{a_1-1}\beta^{b_1-1}\gamma^{c_1-1}\exp[-(a_2\alpha+b_2\beta+c_2\gamma)]}{\Gamma(a_1)\Gamma(b_1)\Gamma(c_1)},$$

  where the mean of the prior is a_1/a_2, b_1/b_2 and c_1/c_2 for α, β and γ, respectively.

- Normal prior distribution

$$g_0(\alpha,\beta,\gamma) \;=\; \left(\frac{a_2b_2c_2}{2\pi}\right)^{1/2}\exp\left\{-\frac{a_2(\alpha-a_1)^2+b_2(\beta-b_1)^2+c_2(\gamma-c_1)^2}{2}\right\},$$

  where the prior mean and variance are a_1 and $1/a_2$ for α, b_1 and $1/b_2$ for β, and c_1 and $1/c_2$ for γ.

- Uniform prior distribution

$$g_0(\alpha,\beta,\gamma) \;=\; \frac{1}{a_2-a_1}\frac{1}{b_2-b_1}\frac{1}{c_2-c_1}I(a_1<\alpha<a_2)I(b_1<\beta<b_2)I(c_1<\gamma<c_2),$$

  where $I(\dot{})$ is an indicator function, and $a_1 < a_2$, $b_1 < b_2$, and $c_1 < c_2$.

For the baseline hazard rate λ, we consider the following scenarios:

- The baseline hazard rate λ is a known constant.
- The prior $g_4(\lambda)$ of the baseline hazard rate λ is distributed as a uniform distribution.
- The prior $g_4(\lambda)$ of the baseline hazard rate λ is distributed as a gamma distribution.

The posterior distributions will be obtained using Markov Chain Monte Carlo (MCMC) sampling method, and the determination of the next dose for the next patient will be based on these posterior distributions as in O'Quigley et al. (1990), Babb et al. (1998), and Chu et al. (2009).

If we assume that the DLT-free survival time for the susceptible patients follows a Weibull distribution and let $g(\alpha, \beta, \gamma, \lambda, v)$ be the prior distribution for α, β, γ, λ and v, the posterior given $\mathcal{H}_n$ is

$$G(\alpha, \beta, \gamma, \lambda, v \mid \mathcal{H}_n) \propto L(\alpha, \beta, \gamma, \lambda, v \mid \mathcal{H}_n) \times g(\alpha, \beta, \gamma, \lambda, v).$$

We assume that the prior distributions of $-\alpha$, β, γ, λ and v are independent such that $g(\alpha, \beta, \gamma, \lambda, v) = g_0(\alpha, \beta, \gamma) g_4(\lambda) g_5(v) = g_1(\alpha) g_2(\beta) g_3(\gamma) g_4(\lambda) g_5(v)$. The priors for $-\alpha$, β and γ are distributed as uniform distributions; the priors for λ and v are distributed as gamma distributions.

3.1.2 CATE Design Based on the Accelerated Failure Time Model

Under the accelerated failure time model,

$$\log(T) \quad = \quad \gamma_0 + \gamma_1 x + \sigma\epsilon, \tag{3.6}$$

where the logarithm of survival time T is linearly related to the covariate dose x with coefficients γ_0 and γ_1. The parameter γ_1 represents the dose effect on time to DLT for patients who will develop a DLT. We assume that dose "accelerates" time to DLT,

which implies that $\gamma_1 < 0$. σ is an unknown scale parameter, and the error term ϵ is a random variable with density function $f(\epsilon)$ and survival function $S(\epsilon)$. The survival time is dependent on both the covariate and the underlying distribution $f(\epsilon)$.

If the error term ϵ follows a double exponential distribution with the following density function $f(\epsilon)$ and survival function $S(\epsilon)$:

$$f(\epsilon) = \exp[\epsilon - \exp(\epsilon)],$$

$$S(\epsilon) = \exp[-\exp(\epsilon)].$$

T has the Weibull distribution with rate λ_x and shape v as

$$\lambda_x = \exp\left(-\frac{\gamma_0 + \gamma_1 x}{\sigma}\right) \text{ and } v = \frac{1}{\sigma}.$$

Let λ_0 be the rate parameter of the weibull distribution for T when dose $x = 0$, then

$$\lambda_0 = \exp\left(-\frac{\gamma_0}{\sigma}\right).$$

T has the following hazard, density and survival functions:

$$h_L(t \mid x, \gamma_0, \gamma_1, \sigma) = \frac{1}{\sigma}\lambda_0 t^{(\frac{1}{\sigma}-1)} \exp\left(-\frac{\gamma_1}{\sigma}x\right),$$

$$S_L(t \mid x, \gamma_0, \gamma_1, \sigma) = \exp\left[-\exp\left(-\frac{\gamma_0 + \gamma_1 x_i}{\sigma}\right) t^{\frac{1}{\sigma}}\right],$$

$$f_L(t \mid x, \gamma_0, \gamma_1, \sigma) = \frac{1}{\sigma}\exp\left(-\frac{\gamma_0 + \gamma_1 x}{\sigma}\right) t^{(\frac{1}{\sigma}-1)} \exp\left[-\exp\left(-\frac{\gamma_0 + \gamma_1 x_i}{\sigma}\right) t^{\frac{1}{\sigma}}\right].$$

The hazard ratio between two individuals i and j is $\exp\left[-\frac{\gamma_1}{\sigma}(x_i - x_j)\right]$ which is not time-dependent. The Weibull AFT model is a special case of the proportional hazards models. The relationship between the parameters $(\gamma_0, \gamma_1, \sigma)$ under the AFT model and

the parameters (γ, λ, v) under the PH model is as follows:

$$\begin{aligned} \gamma_0 &= -\frac{1}{v} log(\lambda), \\ \gamma_1 &= -\frac{1}{v}\gamma, \\ \sigma &= \frac{1}{v}; \end{aligned} \tag{3.7}$$

or

$$\begin{aligned} \gamma &= -\frac{\gamma_1}{\sigma}, \\ \lambda &= \exp\left(-\frac{\gamma_0}{\sigma}\right), \\ v &= \frac{1}{\sigma}. \end{aligned} \tag{3.8}$$

The likelihood (3.4) becomes

$$\begin{aligned} & L(\alpha, \beta, \gamma_0, \gamma_1, \sigma \mid \mathcal{H}_n) \\ &= \prod_{i=1}^{n} [\pi(x_i) f_L(t \mid x_i, \gamma_0, \gamma_1, \sigma)]^{\delta_i} \{\pi(x_i) S_L(t \mid x_i, \gamma_0, \gamma_1, \sigma) + [1 - \pi(x_i)]\}^{1-\delta_i}. \end{aligned}$$

when $\sigma = 1$, T follows the exponential distribution. $\sigma > 0$, and $\gamma_1 < 0$.

In R implementation, the pdf and cdf function for standard gumbel distribution (function rgumbel in VGAM package) is

$$f(w) = \exp[-w - \exp(-w)],$$

$$F(w) = \exp[-\exp(-w)].$$

Therefore, $\epsilon = -w$.

Priors and Posteriors

Assume that the DLT-free survival time for the susceptible patients follows a Weibull distribution. Let $g(\alpha, \beta, \gamma_0, \gamma_1, \sigma)$ be the prior distribution for α, β, γ_0, γ_1 and σ. By

Bayes' theorem the posterior given $\mathcal{H}_n$ is

$$G(\alpha, \beta, \gamma_0, \gamma_1, \sigma \mid \mathcal{H}_n) \propto L(\alpha, \beta, \gamma_0, \gamma_1, \sigma \mid \mathcal{H}_n) \times g(\alpha, \beta, \gamma_0, \gamma_1, \sigma).$$

We assume that the prior distributions of $-\alpha$, β, γ_0, γ_1 and σ are independent such that $g(\alpha, \beta, \gamma_0, \gamma_1, \sigma) = g_1(\alpha)g_2(\beta)g_3(\gamma_0)g_4(\gamma_1)g_5(\sigma)$. The priors for $-\alpha$, β, γ_0 and γ_1 are distributed as uniform distributions; the prior for σ is distributed as inverse-gamma distribution.

3.2 Definition of MTD

The maximum tolerated dose (MTD) can be defined in two ways. First, define MTD_π as the dose at which a target proportion θ of patients will exhibit DLT, which is equivalent to the proportion of susceptible patients:

$$\pi(x = MTD_\pi) = \theta.$$

A second way is to define MTD_T as the dose at which a target proportion θ_τ of patients exhibit DLT during the observation window $[0, \tau]$:

$$\begin{aligned} P(T \leq \tau \mid x = MTD_T) &= 1 - S(\tau \mid x = MTD_T) \\ &= \pi(x = MTD_T)[1 - S_L(\tau \mid x = MTD_T)] \\ &= \theta_\tau, \end{aligned}$$

which is consistent with the traditional definition of MTD.

Both definitions will be evaluated in this thesis. MTD_π is the desired definition, as it is not affected by the planned follow-up time. When the planned follow up time is long enough, MTD_T converges to MTD_π and the MTD dose level under both definitions would be the same. When the planned follow-up time is too short relative to the risk of DLT, the MTD_T definition, equivalent to the traditional definition, will underestimate the true MTD and MTD_π definition is desired.

3.3 Simulation Plan

3.3.1 True Model

Suppose that dose level x_i is the logarithm of the ith dose, $i = 1, ..., k$, and they are set as $x_i = log_{10}(d_i) - log_{10}(d_1)$, for $i = 1, ..., 6$, where $d_{i+1} = 2d_i$. Doses are expressed as multiples of the initial starting dose, and the value of the first dose level will always be zero (i.e., $x_1 = 0$). The standardized dose levels provide robust scheme for treating different scales of the dosage.

Let ρ be the proportion of susceptible patients at the starting does $x = 0$. From (3.2),

$$\rho = \pi(x = 0 \mid \alpha, \beta) = \frac{e^{\alpha}}{1 + e^{\alpha}},$$

then $\alpha = \ln(\rho/(1 - \rho))$. Let $x_{[0.5]}$ be the dose corresponding to a 50% of susceptible patients, then

$$\alpha + \beta x_{[0.5]} = 0 \text{ and } \beta = -\alpha/x_{[0.5]}.$$

Let S_1 be the DLT-free survival probability for the susceptible patients at time τ at the starting dose $x = 0$, and S_2 be the DLT-free survival probability for the susceptible patients at $x_{[0.5]}$.

Exponential Distribution under the Proportional Hazards Model

If the DLT-free survival time for the susceptible patients follows an exponential distribution, from (3.5),

$$S_1 = \exp(-\lambda\tau) \text{ and } S_2 = \exp(-\lambda\tau e^{\gamma x_{[0.5]}}).$$

This implies

$$\lambda = -\ln(S_1)/\tau \text{ and } \gamma = \ln\left(\frac{\ln S_2}{\ln S_1}\right) \Big/ x_{[0.5]}.$$

To check the robustness of different designs to model mis-specification, three true models based on different ρ and $x_{[0.5]}$ are considered. Table 3.1 listed these three dose-response models with low, middle, and high MTD under two different baseline hazards

($\lambda = 0.805$ and 0.403). For all scenarios except Model 3 with $\lambda = 0.403$, MTD_π and MTD_T are the same. In Model 3 with $\lambda = 0.403$, MTD_T (dose level 3) is higher than MTD_π (dose level 2); this is to reflect the shortcoming of insufficiently planned follow-up time τ. When planned follow-up time is short relative to time to develop a DLT, MTD might be over-estimated, and the MTD_π definition is desired. Figure 3.1a shows the dose-response curves by dose based on MTD_π; figure 3.1b and figure 3.1c show the dose-response curves by dose based on two different baseline hazards ($\lambda = 0.805$ and 0.403) and observe only up to time 2 using MTD_T. Note that the definition of MTD_π does not depend on the baseline hazard rate λ. Figure 3.2 and figure 3.3 show the probability of DLT overtime by the pre-defined discrete doses when $\lambda = 0.805$ and $\lambda = 0.403$, respectively.

For these true models, to assess the influence of priors, we consider the following four different independent priors for $-\alpha$, β and γ:

- Uniform $U[0, 10]$ prior distributions for $-\alpha$ and β, and uniform $U[0, 2]$ for γ.
- Exponential distributions with rate = 0.2 for $-\alpha$ and β, and exponential distribution with rate = 1 for γ.
- Normal distributions with mean of 5 and standard deviation of 2.5 for $-\alpha$ and β, and normal distribution with mean of 1 and standard deviation of 0.5 for γ.
- Gamma distributions with shape=4 and rate=0.8 for $-\alpha$ and β, and gamma distribution with shape=4 and rate=4 for γ.

The prior mean for MTD_π is very close to the MTD_π under true model 2; it overestimates MTD_π under true model 3 and underestimates MTD_π under true model 1.

For the baseline hazard rate λ, we consider the following scenarios:

- A known constant ($\lambda = 0.805$ and 0.403) is assumed for λ.
- Uniform $U[0, 2]$ prior distribution is considered for λ when the true $\lambda = 0.805$.

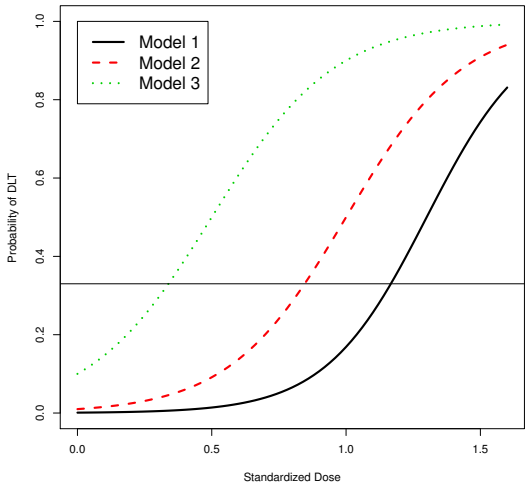
- Gamma prior distribution with shape=56 and rate=70 is considered for λ when the true $\lambda = 0.805$. This prior is chosen so that the probability is around 95% that the true baseline hazard rate falls within (0.6, 1).

- Gamma prior distribution with shape=14 and rate=35 is considered for λ when the true $\lambda = 0.403$. This prior is chosen so that the probability is around 95% that the true baseline hazard rate falls within (0.2, 0.6).

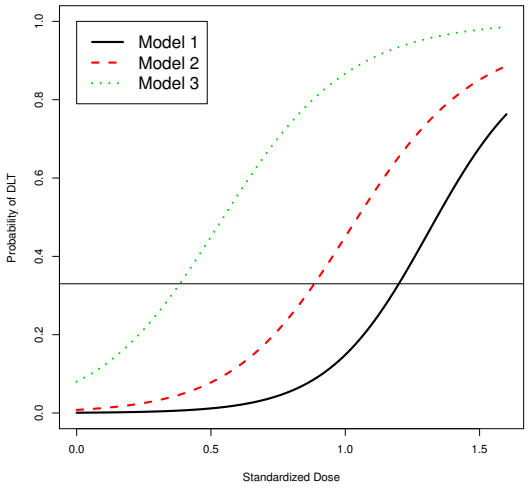
When the true $\lambda = 0.805$, the prior mean for MTD_T is very close to the MTD_T under true model 2; it overestimates MTD_T under true model 3 and underestimates MTD_T under true model 1. When the true $\lambda = 0.403$, the prior mean for MTD_T slightly underestimates the MTD_T under true model 2; it overestimates MTD_T under true model 3 and underestimates MTD_T under true model 1.

Table 3.1: True Models When T Follows Exponential Distribution

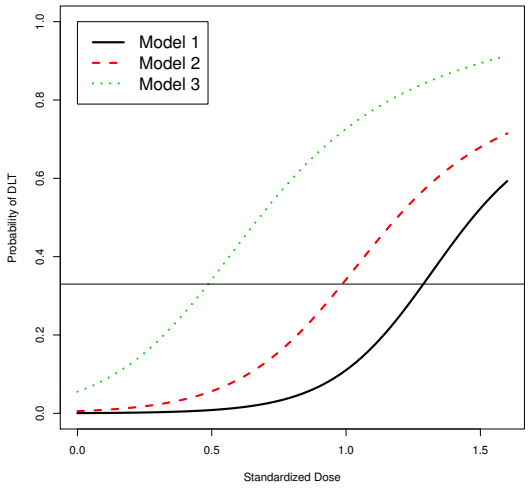
Model	ρ	$x_{[0.5]}$	S_1	S_2	τ	$x_{[0.33]}$	$x^T_{[0.33]}$	MTD	α	β	γ
$\lambda = 0.805$											
1	0.001	1.3	0.20	0.10	2	1.167	1.199	5	-6.907	5.313	0.275
2	0.01	1	0.20	0.10	2	0.846	0.885	4	-4.595	4.595	0.358
3	0.10	0.5	0.20	0.10	2	0.339	0.384	2	-2.197	4.394	0.716
$\lambda = 0.403$											
1	0.001	1.3	0.45	0.32	2	1.167	1.288	5	-6.907	5.313	0.275
2	0.01	1	0.45	0.32	2	0.846	0.986	4	-4.595	4.595	0.358
3	0.10	0.5	0.45	0.32	2	0.339	0.486	2	-2.197	4.394	0.716



a



b



c

Figure 3.1: Three True Models. (a) Probability of susceptible population who will develop a DLT by dose. (b) Probability of DLT with follow-up time=2 by dose when $\lambda = 0.805$. (c) Probability of DLT with follow-up time=2 by dose when $\lambda = 0.403$.

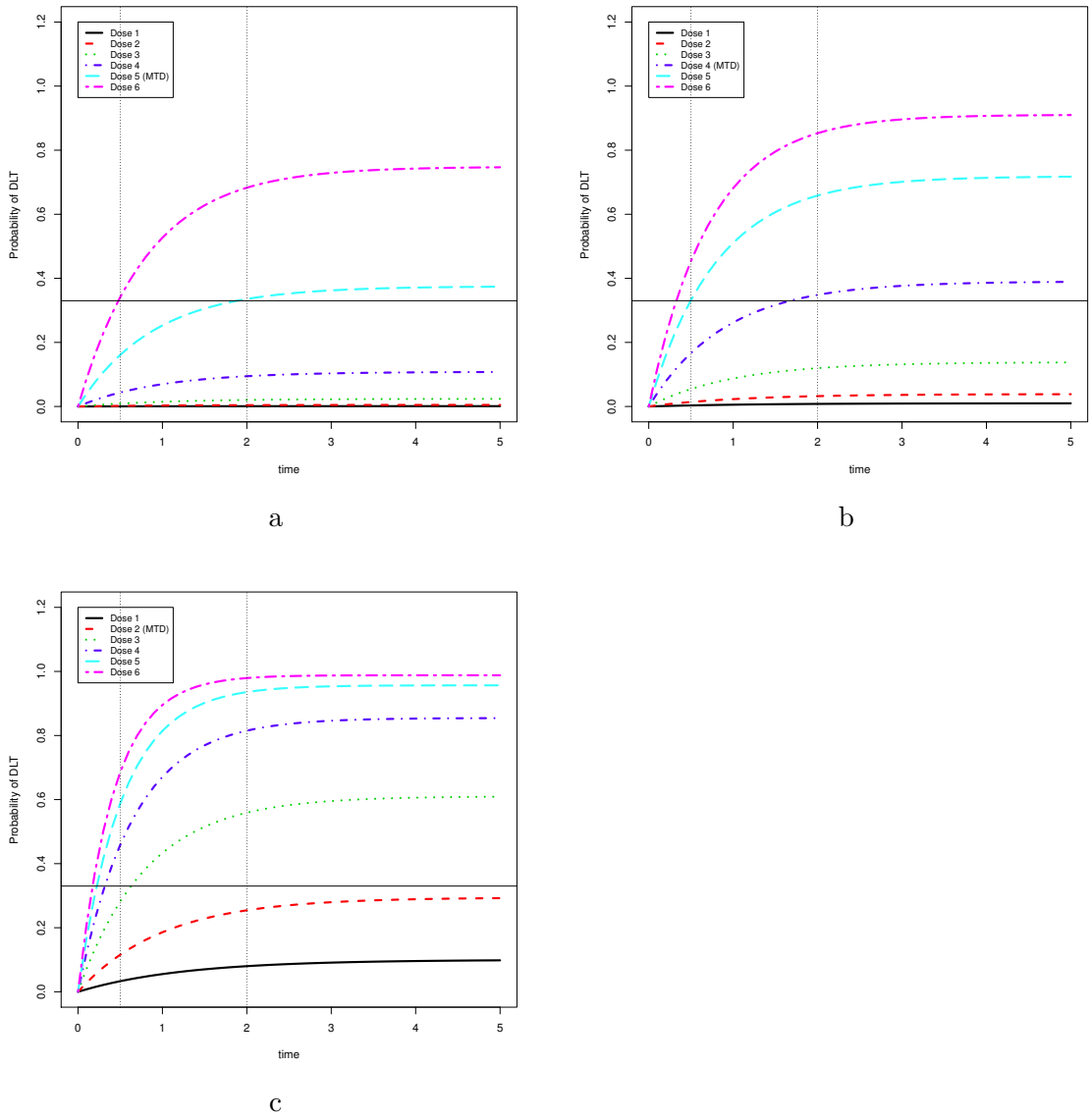


Figure 3.2: Probability of DLT over Time when $\lambda = 0.805$. (a) True Model 1. (b) True Model 2. (c) True Model 3.

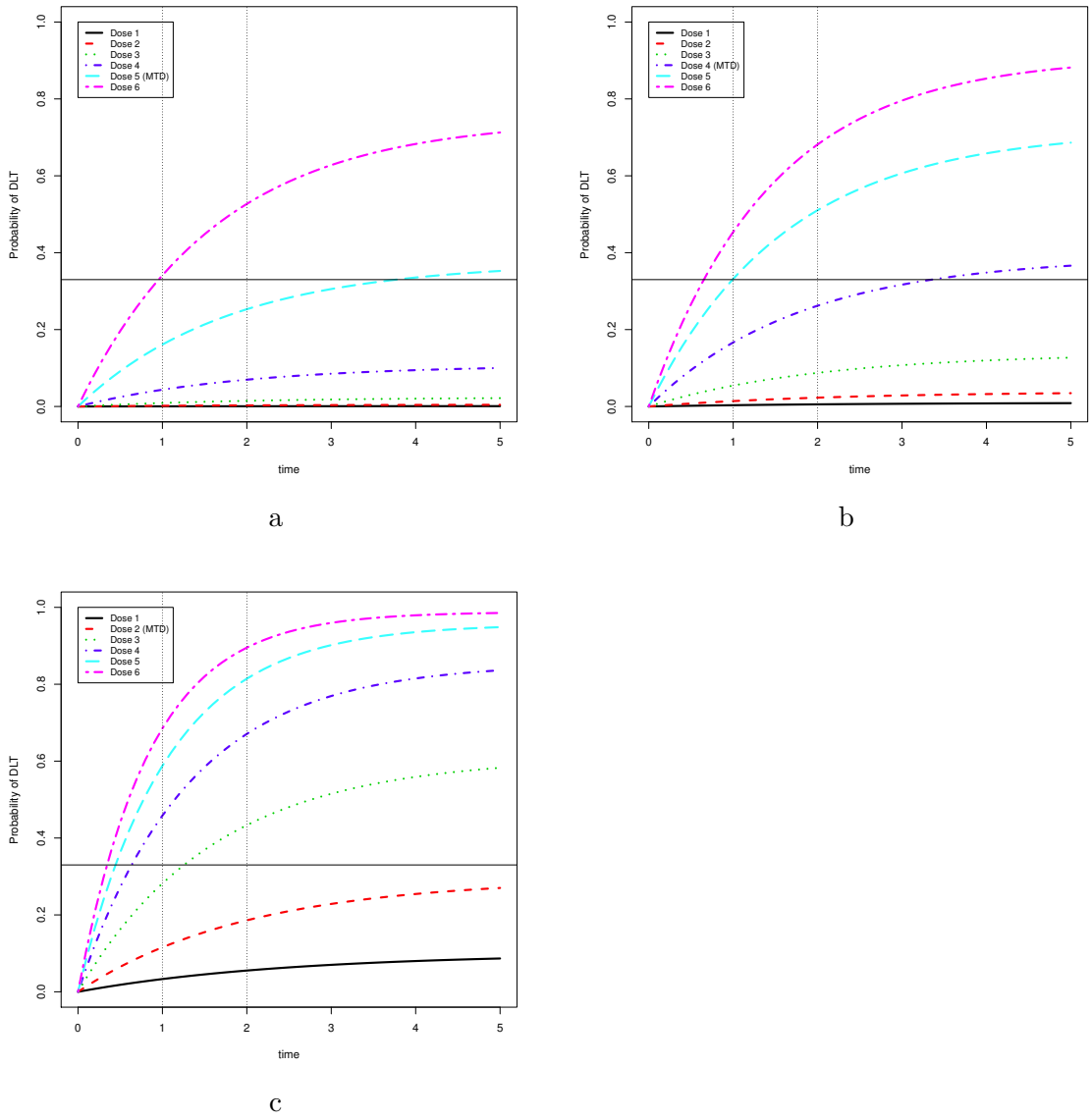


Figure 3.3: Probability of DLT over Time when $\lambda = 0.403$. (a) True Model 1. (b) True Model 2. (c) True Model 3.

Weibull Distribution under the Proportional Hazards Model

Exponential distribution is a special case of Weibull distribution where the shape parameter $v = 1$. We consider the scenario where $v = 2$ and all other parameters (α, β, γ and λ) remain the same true values as defined in Table 3.1.

Table 3.2 listed the three dose-response models with low, middle, and high MTD when $\lambda = 0.805$ and $v = 2$. The following independent priors are considered:

- Uniform $U[0, 10]$ prior distributions for $-\alpha$ and β, and uniform $U[0, 2]$ for γ.
- Gamma distribution with shape=56 and rate=70 for λ.
- Gamma distribution with shape=4 and rate=2 for v.

Table 3.2: True Models When T Follows Weibull Distribution

Model	τ	$x_{[0.33]}$	$x^t_{[0.33]}$	MTD	α	β	γ
$\lambda = 0.805$ and $v = 2$							
1	2	1.167	1.170	5	-6.907	5.313	0.275
2	2	0.846	0.850	4	-4.595	4.595	0.358
3	2	0.339	0.344	2	-2.197	4.394	0.716

Weibull Distribution under the Accelerated Failure Time Model

Table 3.3 listed the three dose-response models considered under the Weibull AFT model; they are equivalent to the true models described above under the proportional hazards model. The following independent priors are considered:

- Uniform $U[0, 10]$ prior distributions for $-\alpha$ and β.
- Uniform $U[0, 0.2]$ prior distributions for γ_0.
- Uniform $U[0, 0.4]$ prior distributions for γ_1.
- Inverse-gamma distribution corresponding to gamma distribution with shape=4 and rate=2 for σ.

Table 3.3: True Models Under Weibull AFT

Model	τ	$x_{[0.33]}$	$x^t_{[0.33]}$	MTD	α	β	γ_0	γ_1
$\sigma = 0.5$								
1	2	1.167	1.170	5	-6.907	5.313	0.108	-0.1375
2	2	0.846	0.850	4	-4.595	4.595	0.108	-0.179
3	2	0.339	0.344	2	-2.197	4.394	0.108	-0.358

3.3.2 Data Generation

Simulate trials of $n = 60$ patients with six dose levels. Assume the target toxicity level $\theta = 0.33$. In each trial, the first patient enters at time 0 and is assigned the starting dose x_1. The planned follow up time τ equals 2 months/cycles. Assume that patients arrive according to the Poisson process with accrual rate of 1. Patients' membership to the population (susceptible or unsusceptible) is first generated following a binomial distribution with mean $\pi(x)$. If the patient belongs to the susceptible population, the underlying time-to-DLT will be generated from the true model for DLT for a susceptible patient. At the time the next patient enters the trial, the observed toxicity indicator (censoring indicator) for all previous patients will be updated and determined by the true time-to-DLT and the current follow-up time.

To avoid a rapid escalation of doses, the next dose cannot be more than one higher dose level of the current highest dose level the patients have been treated. For example, if dose level 2 is the current highest dose, the next dose assignment cannot exceed dose level 3. When dose level 1 is the current dose and if a lower dose is called for, the next assignment would be dose level 1. Similarly, the next assignment will be dose level 6 if the dose higher than dose level 6 is called for.

3.3.3 Data for Comparison to Existing Methods

Data were generated from the true models specified in Chapter 3.3.2. To compare the properties and performance of the CATE design with existing methods (CRM, EWOC, Hybrid, TITE-CRM, TITE-EWOC, and TITE-Hybrid), we consider the following scenarios.

- For conventional CRM, EWOC and Hybrid, a new patient arriving before the evaluation of the last recruited patients have to wait until all previous patients have complete observations.
- A new patient arriving before the evaluation of the last recruited patients will be assigned the dose based on the current complete data.
- Varying the follow-up windows $\tau = 2$ or $\tau = 0.5$.

Two thousand trials are simulated for all the scenarios.

3.3.4 Performance Measure

Let $x_j^{(k)}$ be the dose of jth patient at kth run where $k = 1, ..., K$, where $K(= 2000)$ is the total number of simulation runs. The following measures will be used to compare the different designs (CATE, TITE, and conventional CRM, EWOC, or Hybrid) according to different true models with different prior distributions.

- The convergence rate at jth patient is defined as the proportion of runs that the recommended dose level is the true MTD, i.e.,

$$\frac{\#\{k; x_j^{(k)} = \text{true MTD}\}}{K}.$$

- The overdose proportion at jth patient is defined as the proportion of runs at overdose levels, i.e.

$$\frac{\#\{k; x_j^{(k)} > \text{true MTD}\}}{K}.$$

- Overall trial duration.

Chapter 4

Performance of the CATE Design Based on the Proportional Hazards Model Using Exponential Distribution - Three-Parameter Model

The performance of the CATE Design based on the Proportional Hazards Model using exponential distribution is presented in this chapter. We assume that the baseline hazard rate λ is a known constant. Therefore the CATE design is a an exponential three-parameter models where α, β, γ need to be estimated from data.

4.1 Performance of the Proposed Design

The performance of the proposed design is evaluated under three true dose-response models (low, middle, and high MTD) with high baseline hazard (constant hazard $\lambda = 0.805$) and with planned follow-up time $\tau = 2$. The planned follow-up time of 2 is sufficient for the true dose-response models and the baseline hazard under consideration, where 86%, 89%, and 89% of all DLT on the respective MTD level would have been expected by time 2 under Model 1, 2, and 3, respectively.

4.1.1 Model 1

Based on MTD_π

The percentages of recommendation at each dose level based on MTD_π for various designs, four different priors under the CATE design are summarized in Table 4.1, and the dose assignment proportions are depicted in Figure 4.1 for MTD_π under true model 1 (MTD level = 5). For all the priors examined, all six designs except EWOC have high proportions ($> 83\%$) of recommendation at the correct dose level (level 5). The

EWOC design has lower percentages of correct recommendation compared with other designs, and recommend a lower dose level (level 4) more than the other designs. The hybrid design is comparable to CRM2, CRM3 and CRM4, and slightly better than CRM1. The percentages of recommendation at the overdose level (level 6) are very small ($< 1.8\%$) in all designs. Figure 4.2 shows the convergence rates for MTD_π using the four different prior distributions under true model 1. The convergence rates for the hybrid design are similar to EWOC at the beginning of the trial, and accelerate to similar to the CRM designs with 30 patients. The convergence rates for all designs continued to increase as number of patients increased, but most of the gain had been obtained with 30 patients.

Based on MTD_T

The percentages of recommendation at each dose level for MTD_T under true model 1 (MTD level = 5) are summarized in Table 4.2 and the dose assignment proportions are depicted in Figure 4.3. For all the priors examined, all designs except EWOC have very high proportions ($> 91\%$) of recommendation at the correct dose level (level 5). The performance using MTD_T is better than the corresponding ones using MTD_π. Similar to the cases using MTD_π, the EWOC design has slightly lower percentages of correct recommendation compared with other designs, and recommend a lower dose level (level 4) more often than the other designs. The hybrid design is comparable to CRM2, CRM3 and CRM4, and slightly better than CRM1. The percentages of recommendation at the overdose level (level 6) are very small ($< 1.4\%$) in all designs. Figure 4.4 shows the convergence rates for MTD_T under true model 1. The trend for the convergence rates is similar to those for MTD_π under true model 1. The convergence rates for the hybrid design exceeded EWOC after 20 patients, and are similar to the CRM designs with 30 patients.

Comparison Between MTD_π and MTD_T

Under true model 1 with high MTD (MTD level = 5), the performance using MTD_T is better than the corresponding ones using MTD_π for all six designs and all the priors examined. The relative performance of different designs observed using MTD_T is similar to that using MTD_π. All designs except EWOC have high proportions of recommendation at the correct dose level (level 5). The EWOC design has slightly lower percentages of correct recommendation compared with other designs, and recommend a lower dose level (level 4) more often than the other designs. The performance of the EWOC design using MTD_T is closer to other designs than that using MTD_π. The percentages of recommendation at the overdose level (level 6) are very small in all designs. The trend for the convergence rates is also similar. The convergence rates for the hybrid design exceeded EWOC after 20 patients, and are similar to the CRM designs with 30 patients.

Table 4.1: Percentage of Recommended MTD by Dose Level Under True Model 1 (MTD_π), Exponential Three-Parameter Model

Dose levels	1	2	3	4	5(MTD)	6
Exponential prior						
CRM1	0.00	0.00	0.05	16.35	83.30	0.30
CRM2	0.00	0.00	0.00	7.60	92.05	0.35
CRM3	0.10	0.10	0.65	11.85	85.55	1.75
CRM4	0.00	0.00	0.00	7.95	91.55	0.50
EWOC	0.00	0.00	0.10	23.20	76.65	0.05
Hybrid	0.00	0.00	0.00	8.50	91.05	0.45
Uniform prior						
CRM1	0.00	0.00	0.00	15.80	83.90	0.30
CRM2	0.00	0.00	0.00	7.15	92.75	0.10
CRM3	0.10	0.20	0.25	6.90	91.55	1.00
CRM4	0.00	0.00	0.00	6.70	93.10	0.20
EWOC	0.00	0.00	0.05	24.45	75.40	0.10
Hybrid	0.00	0.00	0.00	7.40	92.15	0.45
Gamma prior						
CRM1	0.00	0.00	0.00	8.00	91.55	0.45
CRM2	0.00	0.00	0.00	3.60	96.05	0.35
CRM3	0.00	0.00	0.00	4.00	95.65	0.35
CRM4	0.00	0.00	0.00	4.80	95.05	0.15
EWOC	0.00	0.00	0.00	16.55	83.25	0.20
Hybrid	0.00	0.00	0.00	4.45	94.95	0.60
Normal prior						
CRM1	0.00	0.00	0.00	12.65	87.15	0.20
CRM2	0.00	0.00	0.00	6.30	93.50	0.20
CRM3	0.00	0.00	0.05	6.70	92.80	0.45
CRM4	0.00	0.00	0.00	5.55	94.25	0.20
EWOC	0.00	0.00	0.00	22.25	77.65	0.10
Hybrid	0.00	0.00	0.00	6.55	93.35	0.10

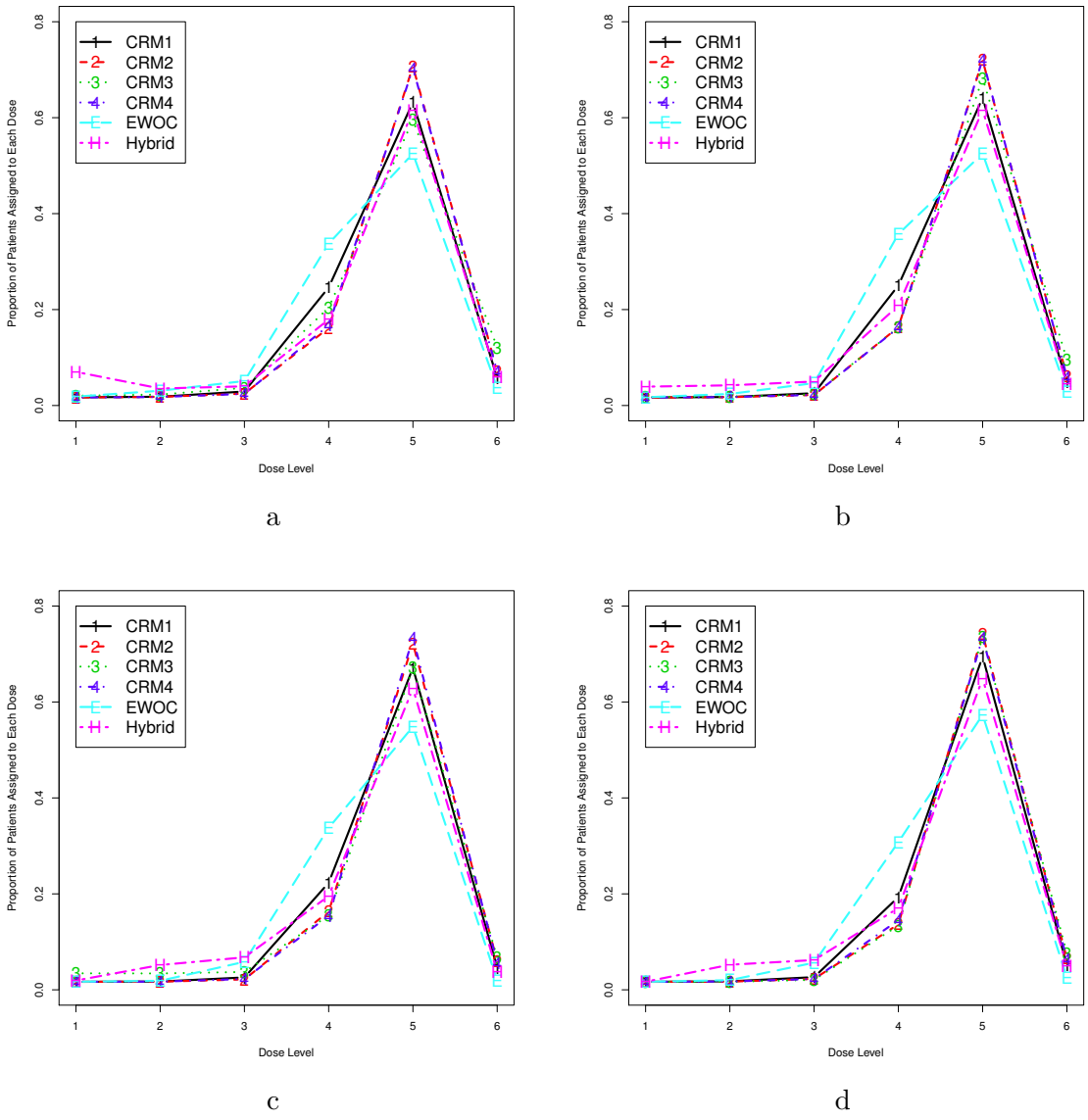


Figure 4.1: Dose assignment proportion with different prior distributions under true Model 1 (MTD_π level = 5). (a) exponential prior, (b) uniform prior, (c) normal prior, and (d) gamma prior.

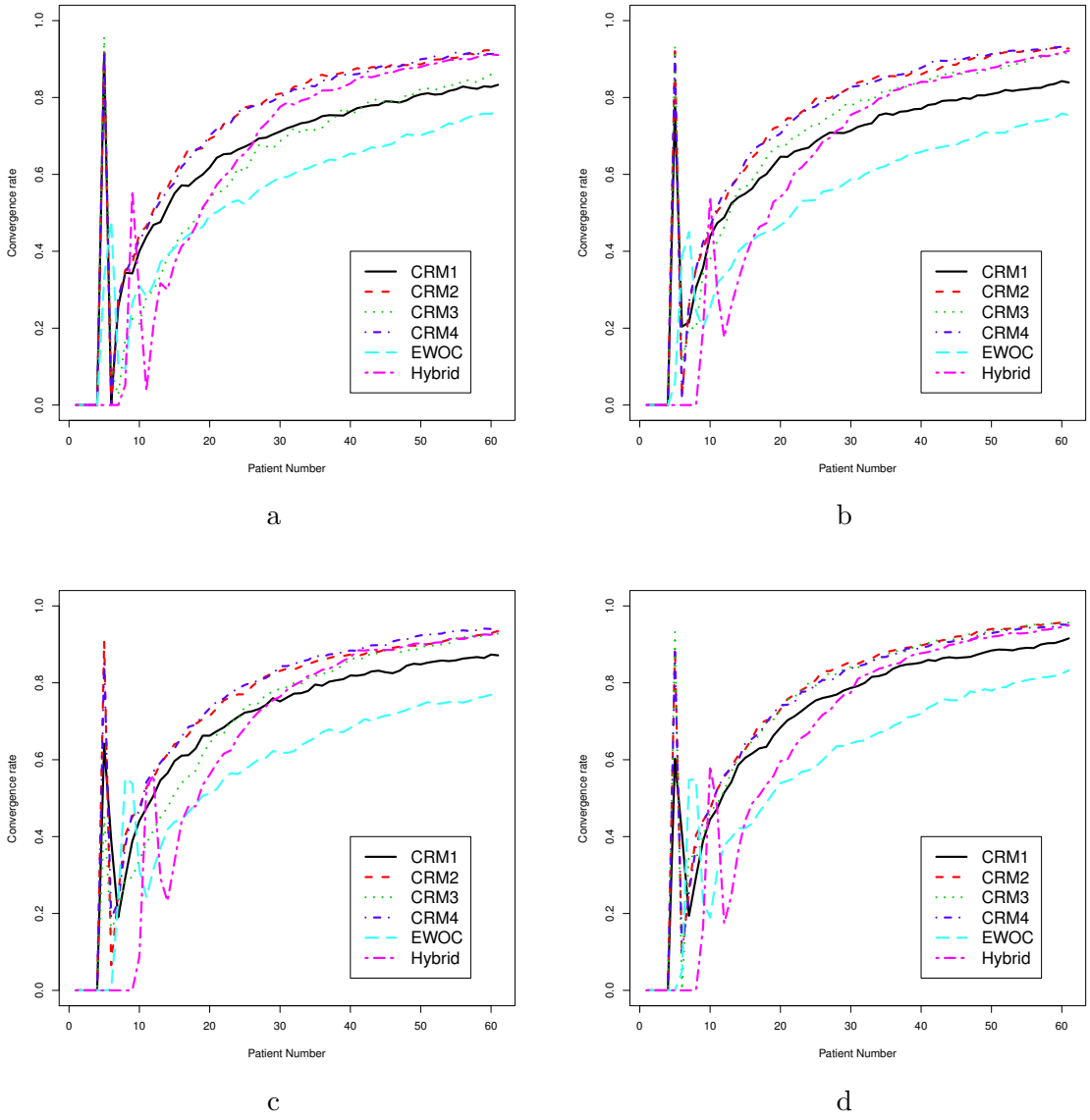


Figure 4.2: Convergence Rates with different prior distributions under true Model 1 (MTD_π level = 5). (a) exponential prior, (b) uniform prior, (c) normal prior, and (d) gamma prior.

Table 4.2: Percentage of Recommended MTD by Dose Level Under True Model 1 (MTD_T), Exponential Three-Parameter Model

Dose levels	1	2	3	4	5(MTD)	6
Exponential prior						
CRM1	0.00	0.00	0.00	8.15	91.40	0.45
CRM2	0.00	0.00	0.00	2.85	95.90	1.25
CRM3	0.00	0.00	0.00	3.10	95.70	1.20
CRM4	0.00	0.00	0.00	3.05	95.65	1.30
EWOC	0.00	0.00	0.00	12.55	87.05	0.40
Hybrid	0.00	0.00	0.00	3.40	95.25	1.35
Uniform prior						
CRM1	0.00	0.00	0.00	7.05	92.75	0.20
CRM2	0.00	0.00	0.00	3.40	96.30	0.30
CRM3	0.00	0.00	0.00	1.90	97.30	0.80
CRM4	0.00	0.00	0.00	2.35	97.25	0.40
EWOC	0.00	0.00	0.00	13.75	86.10	0.15
Hybrid	0.00	0.00	0.00	3.00	96.25	0.75
Gamma prior						
CRM1	0.00	0.00	0.00	4.95	94.80	0.25
CRM2	0.00	0.00	0.00	1.90	97.75	0.35
CRM3	0.00	0.00	0.00	1.60	97.85	0.55
CRM4	0.00	0.00	0.00	2.15	97.45	0.40
EWOC	0.00	0.00	0.00	9.55	90.30	0.15
Hybrid	0.00	0.00	0.00	1.90	97.50	0.60
Normal prior						
CRM1	0.00	0.00	0.00	4.35	95.25	0.40
CRM2	0.00	0.00	0.00	2.20	97.40	0.40
CRM3	0.00	0.00	0.00	1.50	97.90	0.60
CRM4	0.00	0.00	0.00	1.80	97.65	0.55
EWOC	0.00	0.00	0.00	11.40	88.35	0.25
Hybrid	0.00	0.00	0.00	2.20	96.95	0.85

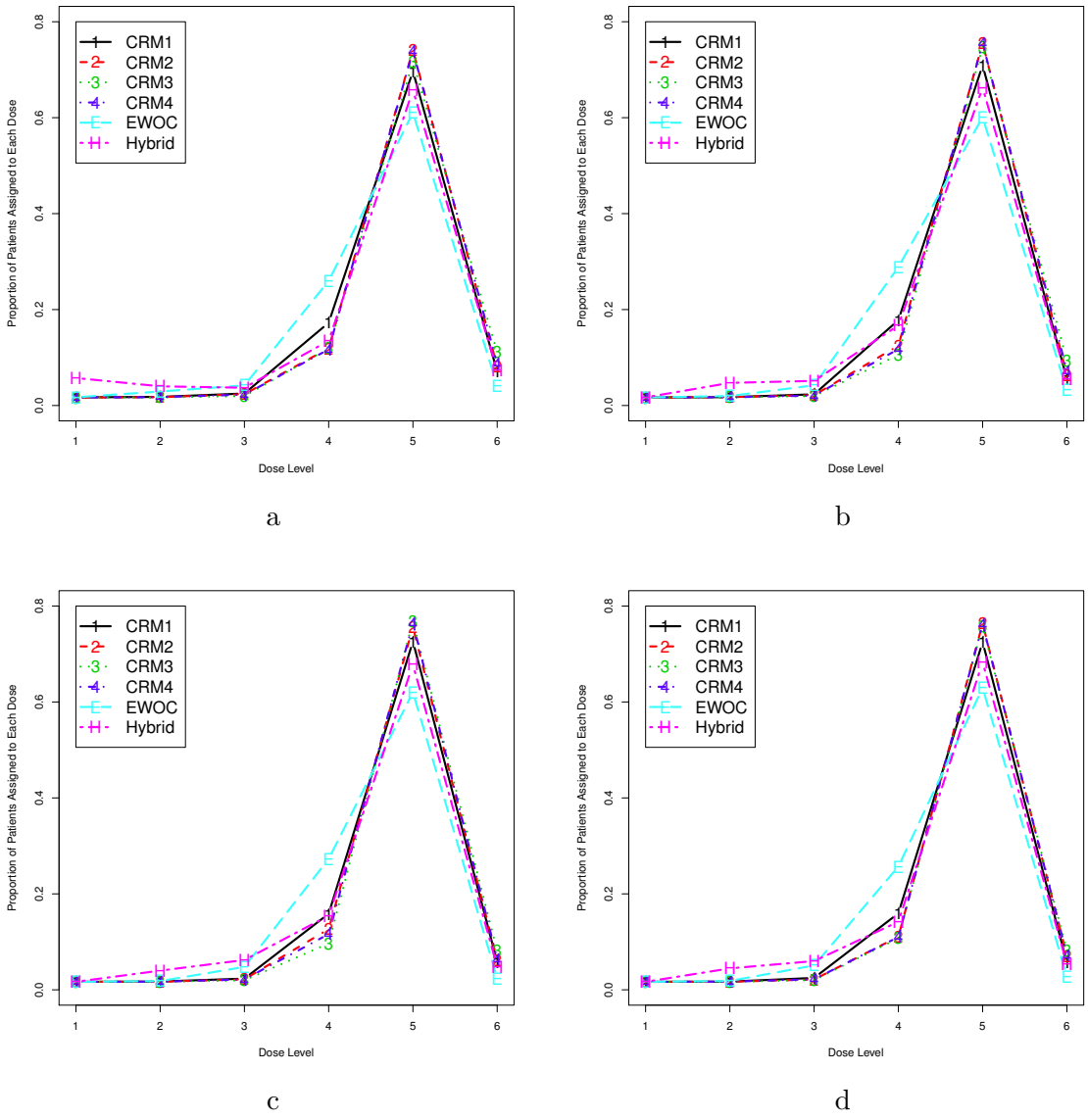


Figure 4.3: Dose assignment proportion with different prior distributions under true Model 1 (MTD_T level = 5). (a) exponential prior, (b) uniform prior, (c) normal prior, and (d) gamma prior.

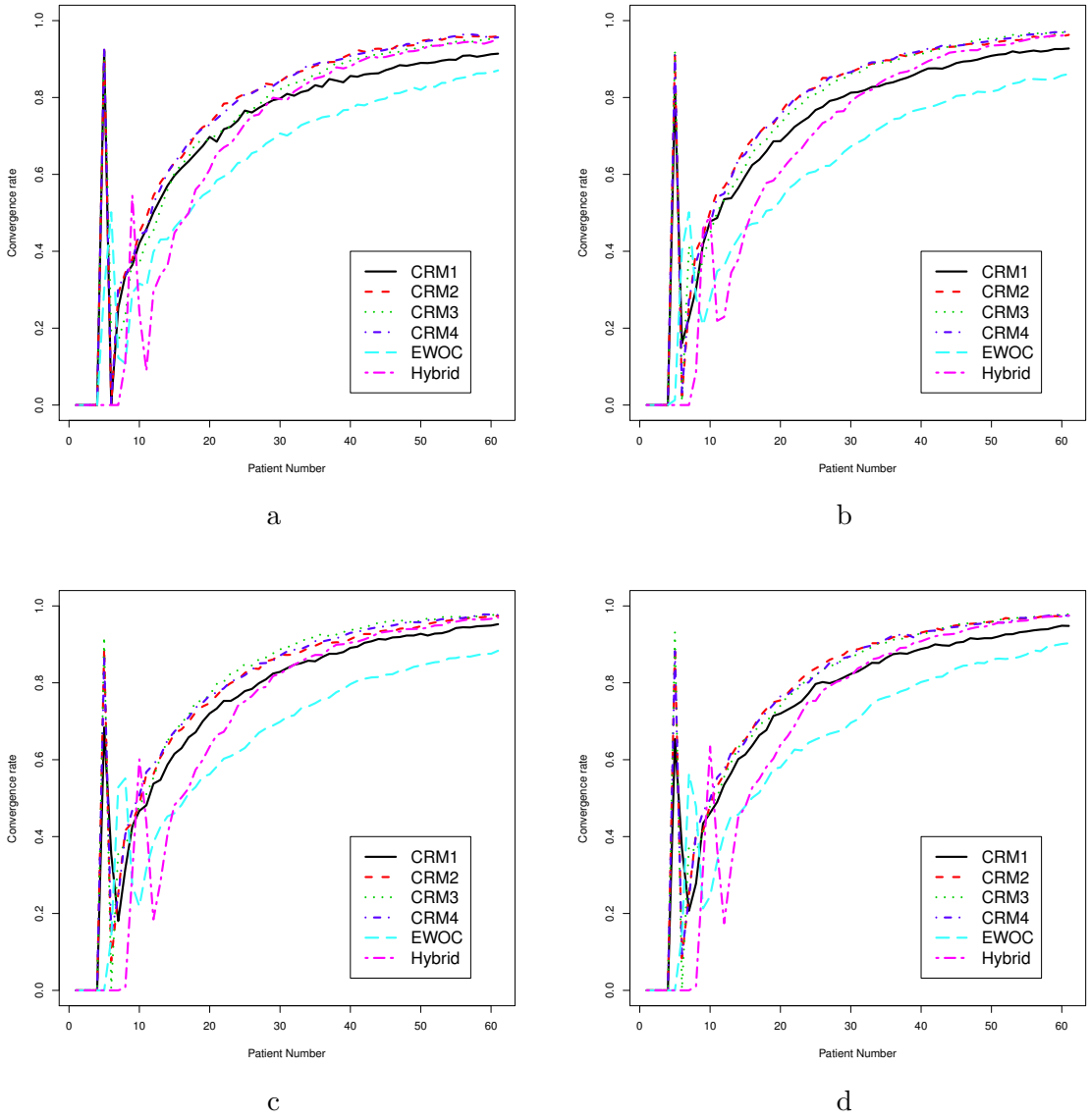


Figure 4.4: Convergence Rates with different prior distributions under true Model 1 (MTD_T level = 5). (a) exponential prior, (b) uniform prior, (c) normal prior, and (d) gamma prior.

4.1.2 Model 2

Based on MTD_π

Table 4.3 shows the percentages of recommendation at each dose level for MTD_π, and Figure 4.5 shows the dose assignment proportions under true model 2 (MTD level = 4). CRM2, CRM3, CRM4, and the hybrid design have comparable high proportions ($> 80\%$) of recommendation at the correct dose level (level 4) for all the different prior distributions considered. The EWOC design have lower percentages of correct recommendation (ranging from 59% to 70%) compared with other designs, and recommend a lower dose level (level 3) more often than the other designs. CRM1 is better than EWOC but lower than other CRM designs and the hybrid design in terms of proportions of correct recommendation. The percentages of recommendation at the overdose levels (level 5 or 6) are very small ($< 1.5\%$) in all designs. As expected, EWOC has the smallest proportions of overdose. Figure 4.6 shows the convergence rates for MTD_π under true model 2. The convergence rates for the hybrid design surpassed EWOC after 20 patients, and surpassed CRM1 after 30 patients and are similar to other CRM designs with 30 patients. The convergence rates for all designs continued to increase as number of patients increased, but most of the gain had been obtained with 30 patients.

Based on MTD_T

Table 4.4 shows the percentages of recommendation at each dose level for MTD_T and Figure 4.7 shows the dose assignment proportions under true model 2 (MTD level = 4). The performance using MTD_T is better than the corresponding ones using MTD_π for all designs and all priors examined. CRM2, CRM3, CRM4, and the hybrid design have comparable high proportions ($> 84\%$) of recommendation at the correct dose level (level 4) for all the different prior distributions considered. Similar to the cases using MTD_π, the EWOC design have lower percentages of correct recommendation compared with other designs, and recommend a lower dose level (level 3) more often than the other designs. The hybrid design is comparable to CRM2, CRM3 and CRM4, and better than CRM1. The percentages of recommendation at the overdose levels (level 5 or 6)

are very small ($< 1.5\%$) in all designs. The trend for the convergence rates is similar to those observed for MTD_π under true model 2 (Figure 4.8).

Comparison Between MTD_π and MTD_T

Under true model 2 with moderate MTD (MTD level = 4), the performance using MTD_T is better than the corresponding ones using MTD_π for all six designs and all the priors examined. The relative performance of different designs observed using MTD_T is similar to that using MTD_π. CRM2, CRM3, CRM4, and the hybrid design have comparable high proportions of recommendation at the correct dose level (level 4). The EWOC design has lower percentages of correct recommendation compared with other designs, and recommends a lower dose level (level 3) more often than the other designs. The performance of the EWOC design using MTD_T is closer to other designs than that using MTD_π. CRM1 is better than EWOC but lower than other CRM designs and the hybrid design. The percentages of recommendation at the overdose levels (level 5 or 6) are very small in all designs. The trend for the convergence rates is also similar. The convergence rates for the hybrid design surpassed EWOC after 20 patients, and surpassed CRM1 after 30 patients and are similar to other CRM designs with 30 patients. The convergence rates for all designs continued to increase as number of patients increased, but most of the gain had been obtained with 30 patients.

Table 4.3: Percentage of Recommended MTD by Dose Level Under True Model 2 (MTD_π), Exponential Three-Parameter Model

Dose levels	1	2	3	4(MTD)	5	6
Exponential prior						
CRM1	0.00	0.45	29.05	70.35	0.15	0.00
CRM2	0.00	0.00	14.45	84.85	0.70	0.00
CRM3	0.25	0.25	17.25	80.75	1.45	0.05
CRM4	0.00	0.00	15.30	84.25	0.45	0.00
EWOC	0.05	0.70	39.65	59.45	0.15	0.00
Hybrid	0.00	0.05	15.05	84.25	0.60	0.05
Uniform prior						
CRM1	0.00	0.00	22.45	77.45	0.10	0.00
CRM2	0.00	0.00	9.80	90.05	0.15	0.00
CRM3	0.05	0.25	12.05	86.95	0.65	0.05
CRM4	0.00	0.00	11.25	88.45	0.30	0.00
EWOC	0.00	0.15	33.05	66.75	0.05	0.00
Hybrid	0.00	0.05	13.30	85.70	0.95	0.00
Gamma prior						
CRM1	0.00	0.00	17.95	81.85	0.20	0.00
CRM2	0.00	0.00	9.35	90.40	0.25	0.00
CRM3	0.00	0.00	9.35	90.20	0.45	0.00
CRM4	0.00	0.00	8.25	91.25	0.50	0.00
EWOC	0.00	0.05	30.00	69.95	0.00	0.00
Hybrid	0.00	0.00	9.75	89.65	0.60	0.00
Normal prior						
CRM1	0.00	0.00	21.10	78.85	0.05	0.00
CRM2	0.00	0.00	10.35	89.60	0.05	0.00
CRM3	0.00	0.00	9.55	89.70	0.75	0.00
CRM4	0.00	0.00	9.00	90.95	0.05	0.00
EWOC	0.00	0.05	34.10	65.85	0.00	0.00
Hybrid	0.00	0.00	11.30	88.10	0.60	0.00

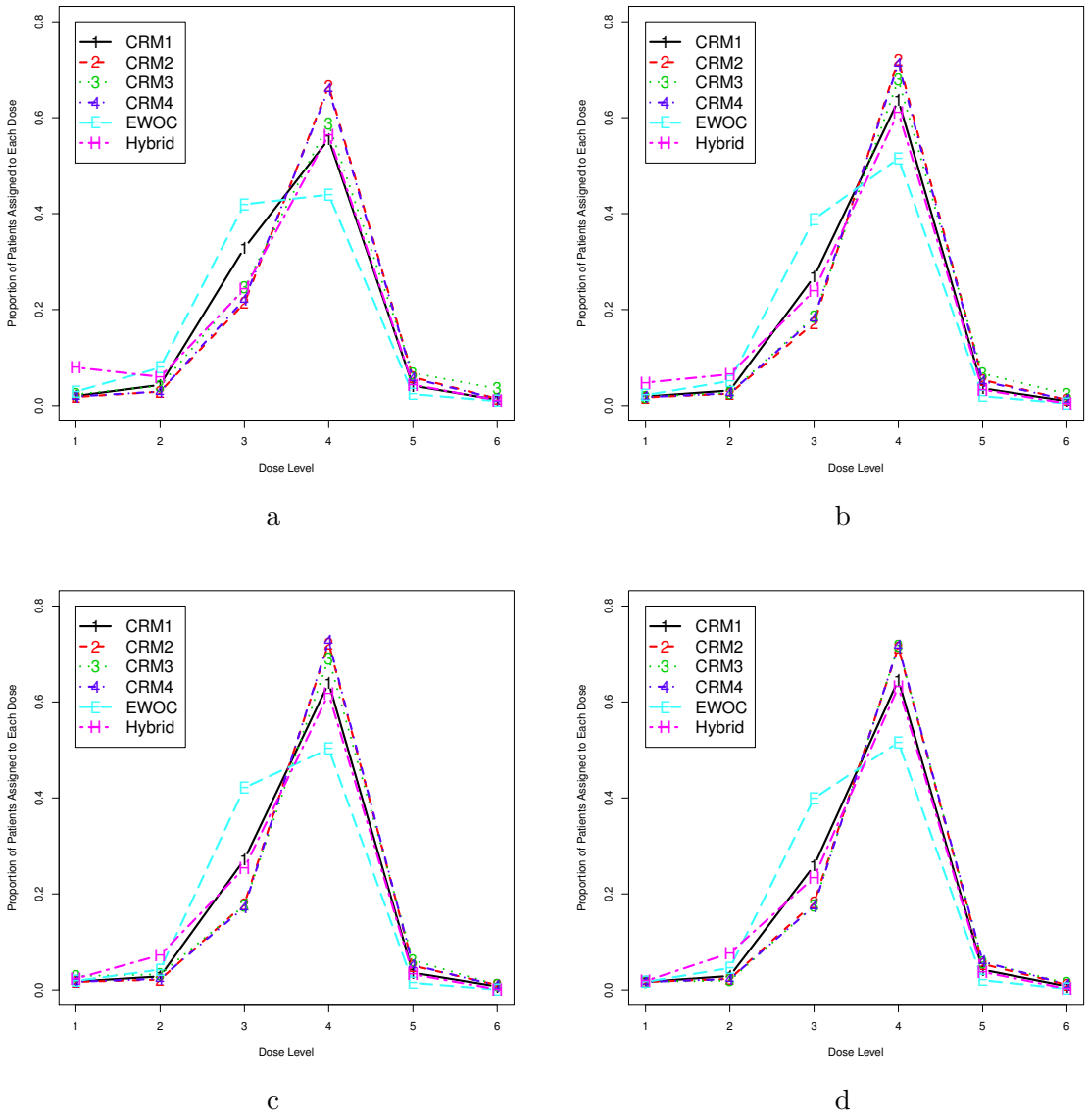


Figure 4.5: Dose assignment proportion with different prior distributions under true Model 2 (MTD_π level = 4). (a) exponential prior, (b) uniform prior, (c) normal prior, and (d) gamma prior.

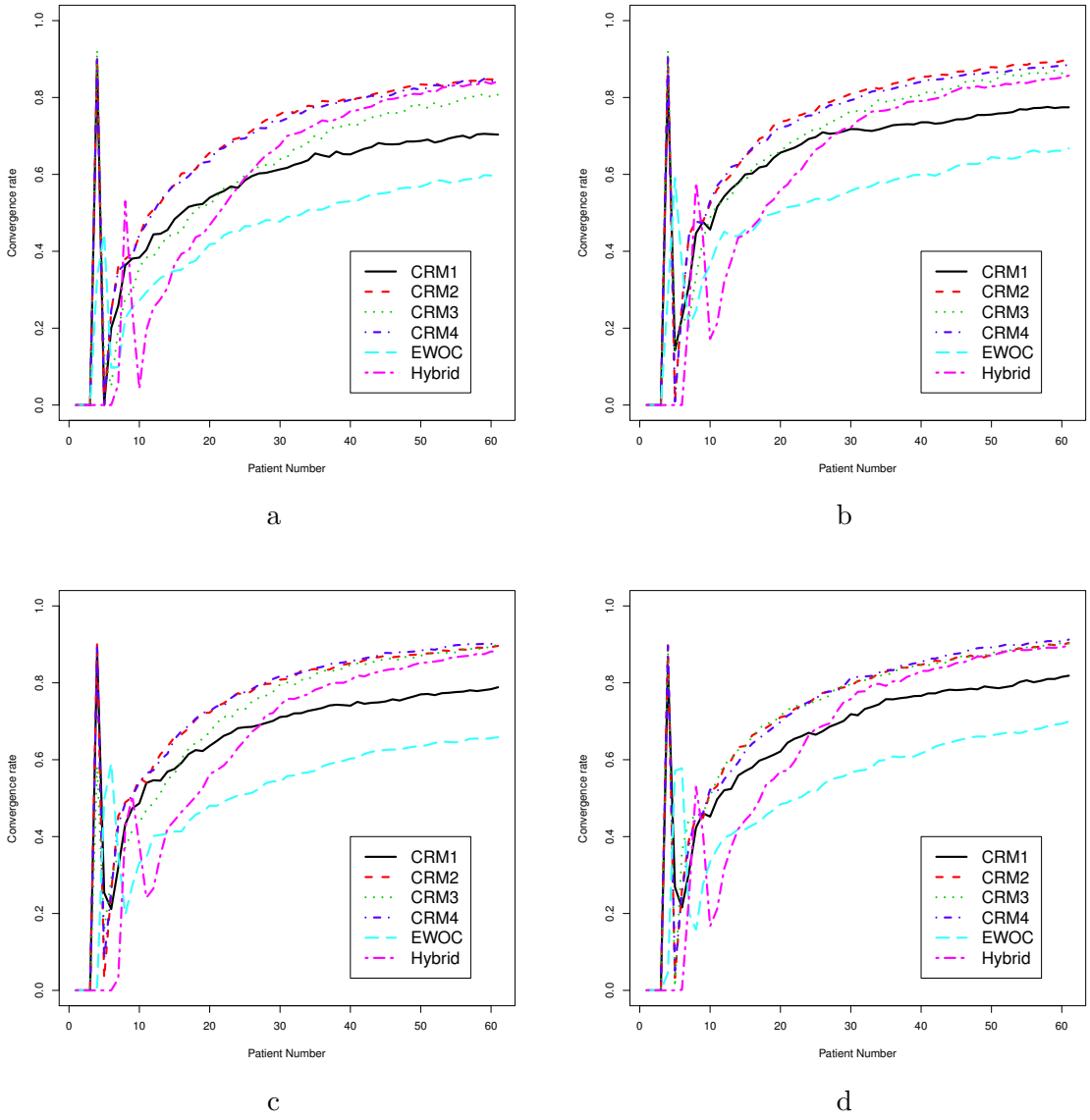


Figure 4.6: Convergence Rates with different prior distributions under true Model 2 (MTD_π level = 4). (a) exponential prior, (b) uniform prior, (c) normal prior, and (d) gamma prior.

Table 4.4: Percentage of Recommended MTD by Dose Level Under True Model 2 (MTD_T), Exponential Three-Parameter Model

Dose levels	1	2	3	4(MTD)	5	6
Exponential prior						
CRM1	0.00	0.05	15.05	84.35	0.55	0.00
CRM2	0.00	0.00	7.45	91.35	1.20	0.00
CRM3	0.00	0.00	5.10	93.90	1.00	0.00
CRM4	0.00	0.00	6.30	92.35	1.35	0.00
EWOC	0.00	0.40	23.90	75.55	0.15	0.00
Hybrid	0.00	0.00	7.35	91.60	1.05	0.00
Uniform prior						
CRM1	0.00	0.00	12.20	87.25	0.55	0.00
CRM2	0.00	0.00	5.65	93.65	0.70	0.00
CRM3	0.00	0.00	3.50	95.70	0.80	0.00
CRM4	0.00	0.00	4.55	94.70	0.75	0.05
EWOC	0.00	0.15	21.05	78.65	0.15	0.00
Hybrid	0.00	0.00	7.45	91.10	1.45	0.00
Gamma prior						
CRM1	0.00	0.00	10.45	89.15	0.40	0.00
CRM2	0.00	0.00	4.75	94.85	0.40	0.00
CRM3	0.00	0.00	3.05	96.50	0.45	0.00
CRM4	0.00	0.00	4.65	94.75	0.60	0.00
EWOC	0.00	0.00	19.70	80.10	0.20	0.00
Hybrid	0.00	0.00	5.65	93.65	0.70	0.00
Normal prior						
CRM1	0.00	0.00	9.60	90.05	0.30	0.05
CRM2	0.00	0.00	5.60	93.90	0.50	0.00
CRM3	0.00	0.00	3.65	95.50	0.85	0.00
CRM4	0.00	0.00	5.10	94.20	0.70	0.00
EWOC	0.00	0.05	21.60	78.35	0.00	0.00
Hybrid	0.00	0.05	6.10	93.25	0.65	0.00

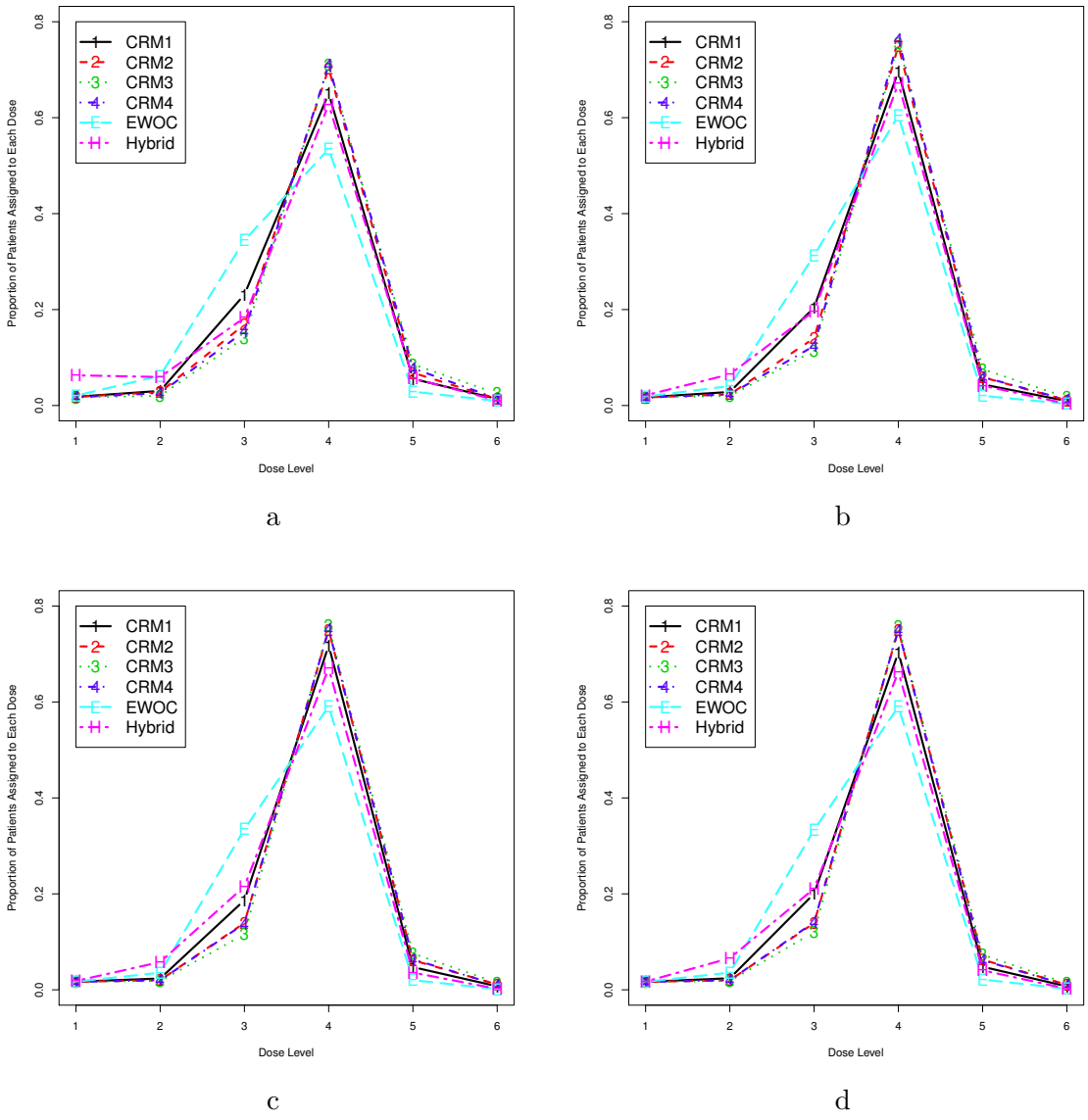


Figure 4.7: Dose assignment proportion with different prior distributions under true Model 2 (MTD_T level = 4). (a) exponential prior, (b) uniform prior, (c) normal prior, and (d) gamma prior.

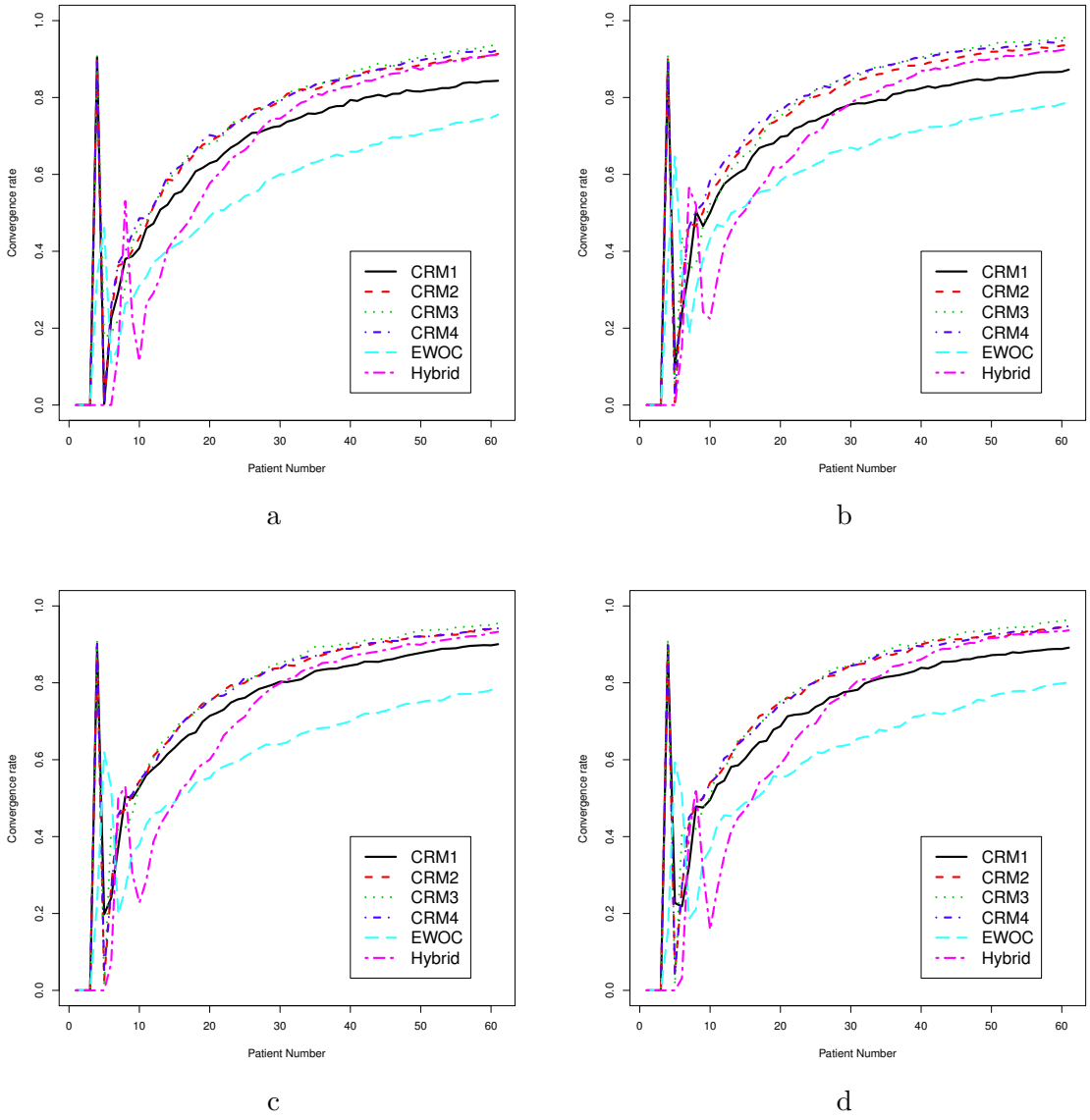


Figure 4.8: Convergence Rates with different prior distributions under true Model 2 (MTD_T level = 4) (a) exponential prior, (b) uniform prior, (c) normal prior, and (d) gamma prior.

4.1.3 Model 3

Based on MTD_π

The percentages of recommendation at each dose level for MTD_π under true model 3 (MTD level = 2) are summarized in Table 4.5, and the dose assignment proportions are depicted in Figure 4.9. All designs except EWOC with exponential prior have high proportions ($> 83\%$) of recommendation at the correct dose level (level 2). The EWOC design with exponential prior has lower percentage of correct recommendation, and recommends a lower dose level (level 1) more often than other designs. When uniform prior distributions are used, CRM1 design has slightly higher proportion of recommendation at the correct dose level (level 2) than the other designs. When gamma prior or normal prior are used, the EWOC design has the comparable performance as the CRM1 design, and is slightly better than the other designs in terms of proportions of correct recommendation. As expected, EWOC has the smallest proportions of recommendation at the overdose level (level 3, 4, 5 or 6) for all the prior distributions examined. The hybrid design recommends fewer overdose levels than CRM2, CRM3, and CRM4 do when gamma prior or normal prior are used, and is comparable to CRM2, CRM3, or CRM4 when exponential prior or uniform prior are used. The convergence rates for all designs increase as number of patients increased in the trial as shown in Figure 4.10. When exponential priors are used, the convergence rate for EWOC is lower than other designs. The hybrid design exceeded EWOC after 20 patients and is comparable to the CRM designs. When uniform priors are used, the convergence rate for CRM1 design is higher than the other designs, and the hybrid design are slightly better than the EWOC designs with 20 patients. The EWOC design and the CRM1 design have similar convergence rates and are higher than the other designs when normal or gamma priors are considered.

Based on MTD_T

Table 4.6 shows the percentages of recommendation at each dose level for MTD_T under true model 3 (MTD level = 2), and Figure 4.11 shows the dose assignment proportions.

CRM2 design under the normal prior failed to converge and is not listed in the table. The EWOC design has higher proportion of correct recommendation using MTD_T than the corresponding ones using MTD_π in all priors examined except the gamma prior; whereas the performance for other designs using MTD_T is worse than the corresponding ones using MTD_π for all priors examined except for CRM1 under exponential prior. This is likely due to the underlying dose-toxicity curve as shown in Figure 3.2 (c). The EWOC design has higher proportion of recommendation at the correct dose level (level 2) than the other designs, and has the smallest recommendation at the overdose level for all the prior distributions examined. The hybrid design recommends fewer overdose levels than CRM2, CRM3, and CRM4 do except for one case where CRM2 has slightly fewer overdose than the hybrid design under uniform prior. The EWOC design has the highest convergence rates followed by the CRM1 design under uniform, gamma or normal prior (Figure 4.12). When exponential prior is considered, the CRM1 has the highest convergence rate followed by EWOC. The hybrid design has similar convergence rates with EWOC at the beginning of the trial, but stay at about the same rates with 30 patients entered the trial; while the convergence rates for EWOC and CRM1 design continue to increase as number of patients increased.

Comparison Between MTD_π and MTD_T

Under true model 3 with low MTD (MTD level = 2), the EWOC design has higher proportion of correct recommendation using MTD_T than the corresponding ones using MTD_π in all priors examined except the gamma prior; whereas the performance for other designs using MTD_T is worse than the corresponding ones using MTD_π for all priors examined except for CRM1 under exponential prior. The proportion of recommendation at the overdose level is higher using MTD_T than the corresponding ones using MTD_π. This is likely due to the underlying dose-toxicity curve. As shown in Figure 3.2 (c), the probability of DLT under true model 3 at dose level 2 is closer to the target toxicity level θ for the MTD_π definition (at time 5) than for the MTD_T definition (at time 2). When exponential priors are used, the convergence rate for EWOC is lower than other designs using MTD_π, while using MTD_T the CRM1 has the highest

convergence rate followed by EWOC. When uniform priors are used, the convergence rate for CRM1 design is the highest and that for EWOC design is the lowest using MTD_π. The EWOC design and the CRM1 design have similar convergence rates and are higher than the other designs when normal or gamma priors are considered using MTD_π. When using MTD_T, the EWOC design has the highest convergence rates followed by the CRM1 design under uniform, gamma or normal prior.

Table 4.5: Percentage of Recommended MTD by Dose Level Under True Model 3 (MTD_π), Exponential Three-Parameter Model

Dose levels	1	2(MTD)	3	4	5	6
Exponential prior						
CRM1	13.65	83.45	2.90	0.00	0.00	0.00
CRM2	4.90	85.20	9.90	0.00	0.00	0.00
CRM3	5.60	85.75	8.40	0.25	0.00	0.00
CRM4	5.10	86.70	8.20	0.00	0.00	0.00
EWOC	19.70	78.00	2.25	0.05	0.00	0.00
Hybrid	5.80	84.55	9.65	0.00	0.00	0.00
Uniform prior						
CRM1	5.55	91.05	3.40	0.00	0.00	0.00
CRM2	1.55	88.50	9.95	0.00	0.00	0.00
CRM3	3.65	88.00	8.30	0.05	0.00	0.00
CRM4	1.50	86.65	11.85	0.00	0.00	0.00
EWOC	11.05	86.75	2.20	0.00	0.00	0.00
Hybrid	2.60	87.95	9.40	0.05	0.00	0.00
Gamma prior						
CRM1	1.55	94.95	3.50	0.00	0.00	0.00
CRM2	0.15	89.30	10.55	0.00	0.00	0.00
CRM3	0.15	90.00	9.85	0.00	0.00	0.00
CRM4	0.15	89.45	10.40	0.00	0.00	0.00
EWOC	3.35	95.05	1.60	0.00	0.00	0.00
Hybrid	0.80	90.50	8.70	0.00	0.00	0.00
Normal prior						
CRM1	2.75	93.40	3.85	0.00	0.00	0.00
CRM2	0.65	86.70	12.65	0.00	0.00	0.00
CRM3	1.20	88.35	10.45	0.00	0.00	0.00
CRM4	0.45	86.25	13.30	0.00	0.00	0.00
EWOC	5.40	92.15	2.45	0.00	0.00	0.00
Hybrid	1.30	89.45	9.25	0.00	0.00	0.00

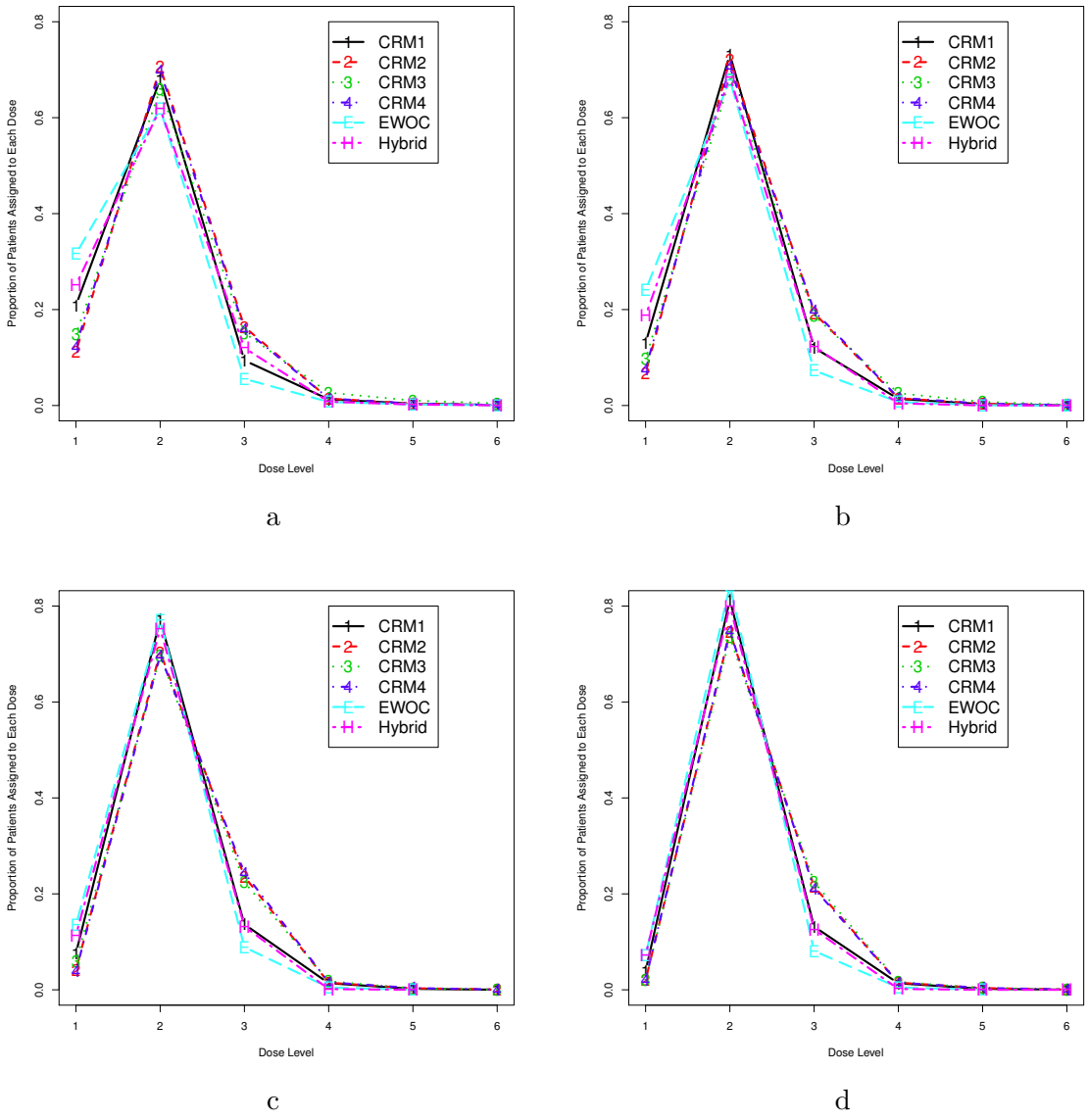


Figure 4.9: Dose assignment proportion with different prior distributions under true Model 3 (MTD_π level =2). (a) exponential prior, (b) uniform prior, (c) normal prior, and (d) gamma prior.

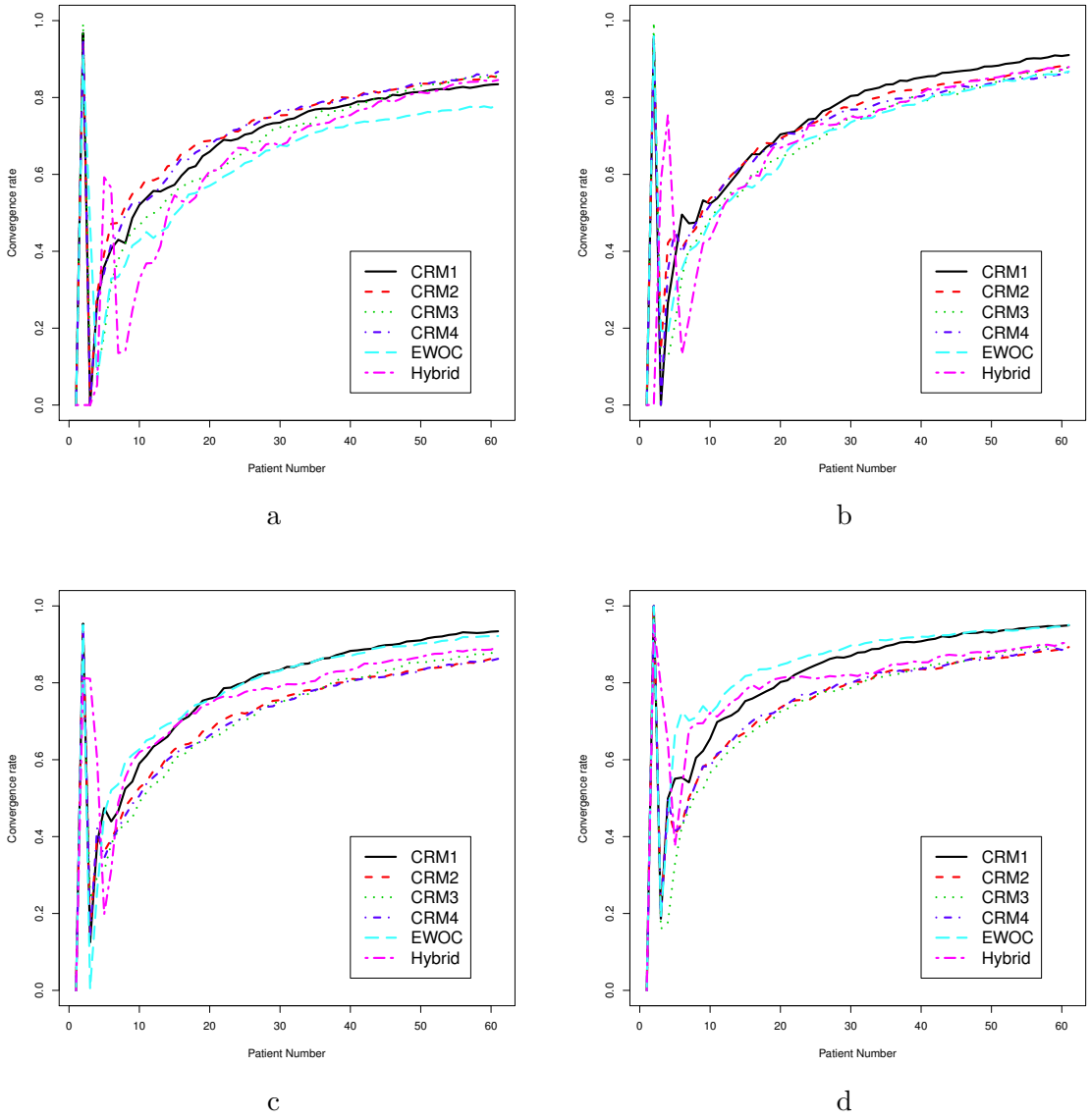


Figure 4.10: Convergence Rates with different prior distributions under true Model 3 (MTD_π level = 2) (a) exponential prior, (b) uniform prior, (c) normal prior, and (d) gamma prior.

Table 4.6: Percentage of Recommended MTD by Dose Level Under True Model 3 (MTD_T), Exponential Three-Parameter Model

Dose levels	1	2(MTD)	3	4	5	6
Exponential prior						
CRM1	1.00	90.75	8.25	0.00	0.00	0.00
CRM2	0.40	78.05	21.55	0.00	0.00	0.00
CRM3	0.15	80.60	19.25	0.00	0.00	0.00
CRM4	0.45	78.75	20.80	0.00	0.00	0.00
EWOC	3.00	91.70	5.30	0.00	0.00	0.00
Hybrid	0.75	82.65	16.60	0.00	0.00	0.00
Uniform prior						
CRM1	1.95	88.85	9.20	0.00	0.00	0.00
CRM2	0.40	80.55	19.05	0.00	0.00	0.00
CRM3	0.05	80.30	19.65	0.00	0.00	0.00
CRM4	0.35	77.20	22.45	0.00	0.00	0.00
EWOC	3.30	91.75	4.95	0.00	0.00	0.00
Hybrid	0.45	80.05	19.40	0.10	0.00	0.00
Gamma prior						
CRM1	0.15	90.55	9.30	0.00	0.00	0.00
CRM2	0.05	76.65	23.30	0.00	0.00	0.00
CRM3	0.00	78.55	21.45	0.00	0.00	0.00
CRM4	0.00	77.70	22.30	0.00	0.00	0.00
EWOC	0.65	94.30	5.00	0.05	0.00	0.00
Hybrid	0.05	81.30	18.65	0.00	0.00	0.00
Normal prior						
CRM1	0.40	89.60	10.00	0.00	0.00	0.00
CRM2						
CRM3	0.00	77.15	22.85	0.00	0.00	0.00
CRM4	0.00	73.55	26.45	0.00	0.00	0.00
EWOC	0.65	93.60	5.75	0.00	0.00	0.00
Hybrid	0.15	81.55	18.30	0.00	0.00	0.00

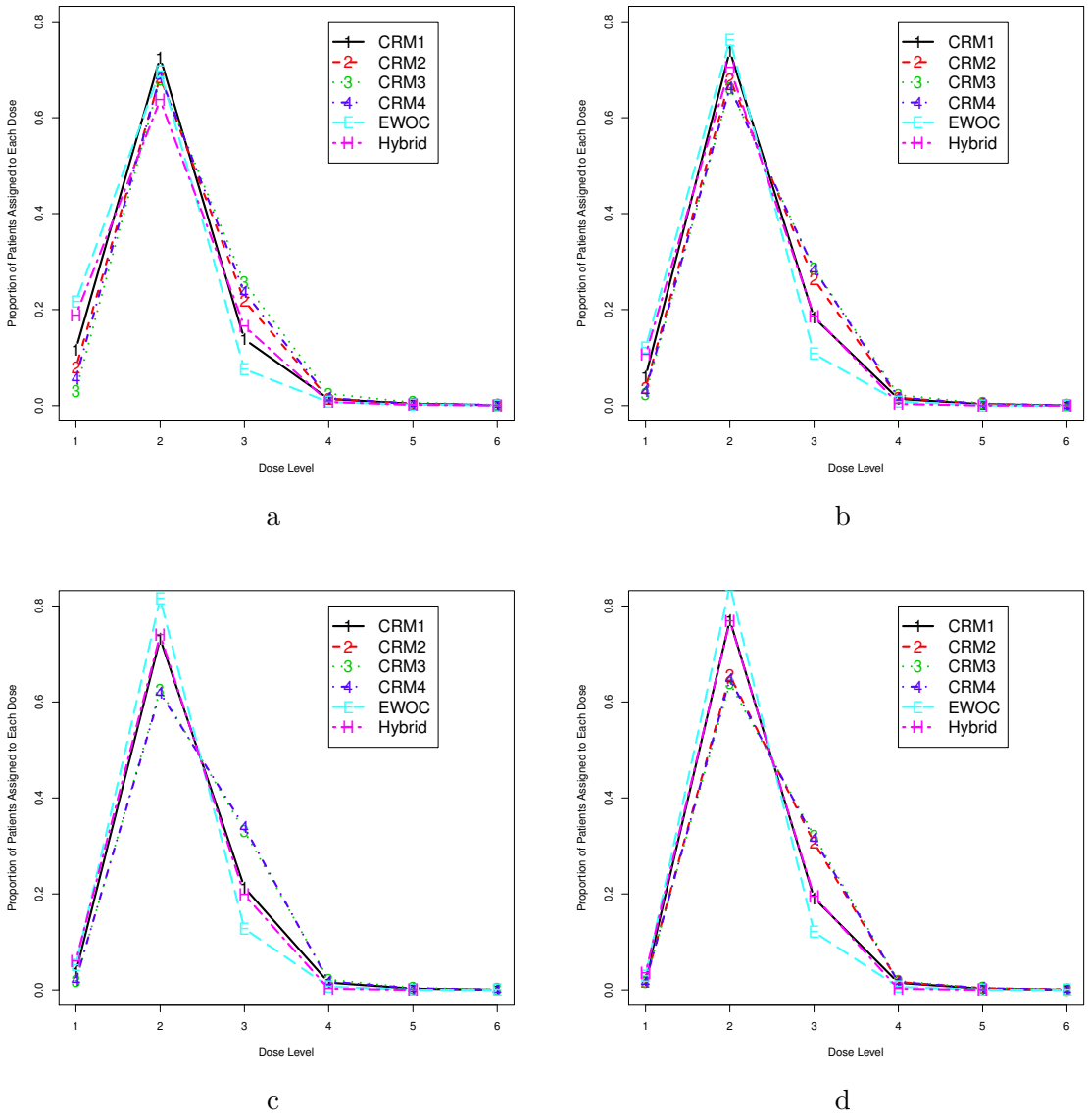


Figure 4.11: Dose assignment proportion with different prior distributions under true Model 3 (MTD_T level =2). (a) exponential prior, (b) uniform prior, (c) normal prior, and (d) gamma prior.

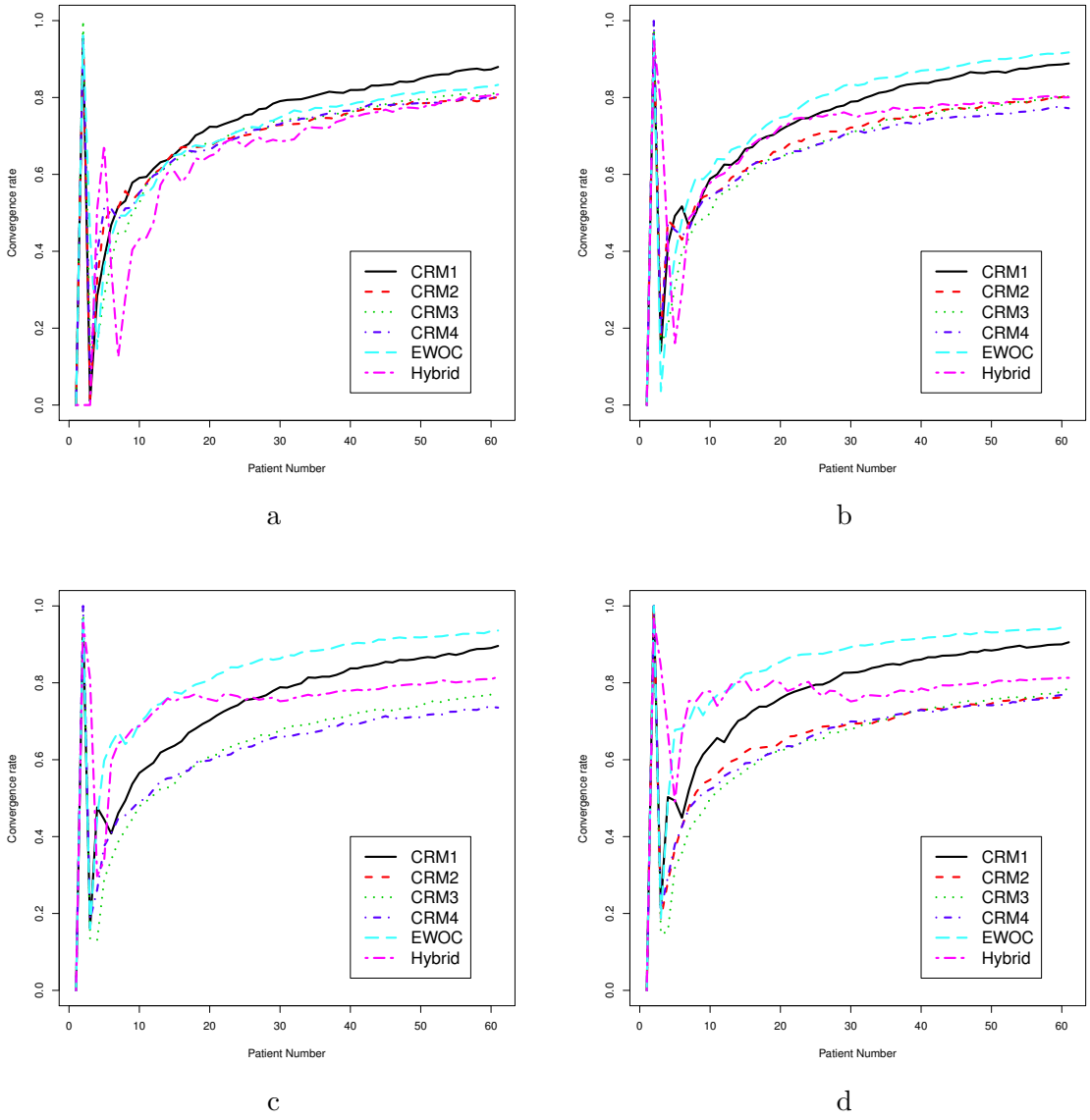


Figure 4.12: Convergence Rates with different prior distributions under true Model 3 (MTD_T level = 2) (a) exponential prior, (b) uniform prior, (c) normal prior, and (d) gamma prior.

4.1.4 Overdose Comparisons Under Different True Models and Different Prior Distributions

The percentage of recommendation at the overdose level for each design under different true models and different prior distributions is depicted in Figure 4.13. EWOC design has the lowest percentages in all scenarios except under true model 1 (MTD_π). The percentages of overdose for all designs are very low under true model 1 or true model 2 ($< 1\%$ for most designs). Under true model 3, the hybrid design recommends fewer overdose levels than CRM2, CRM3, and CRM4 do in most scenarios.

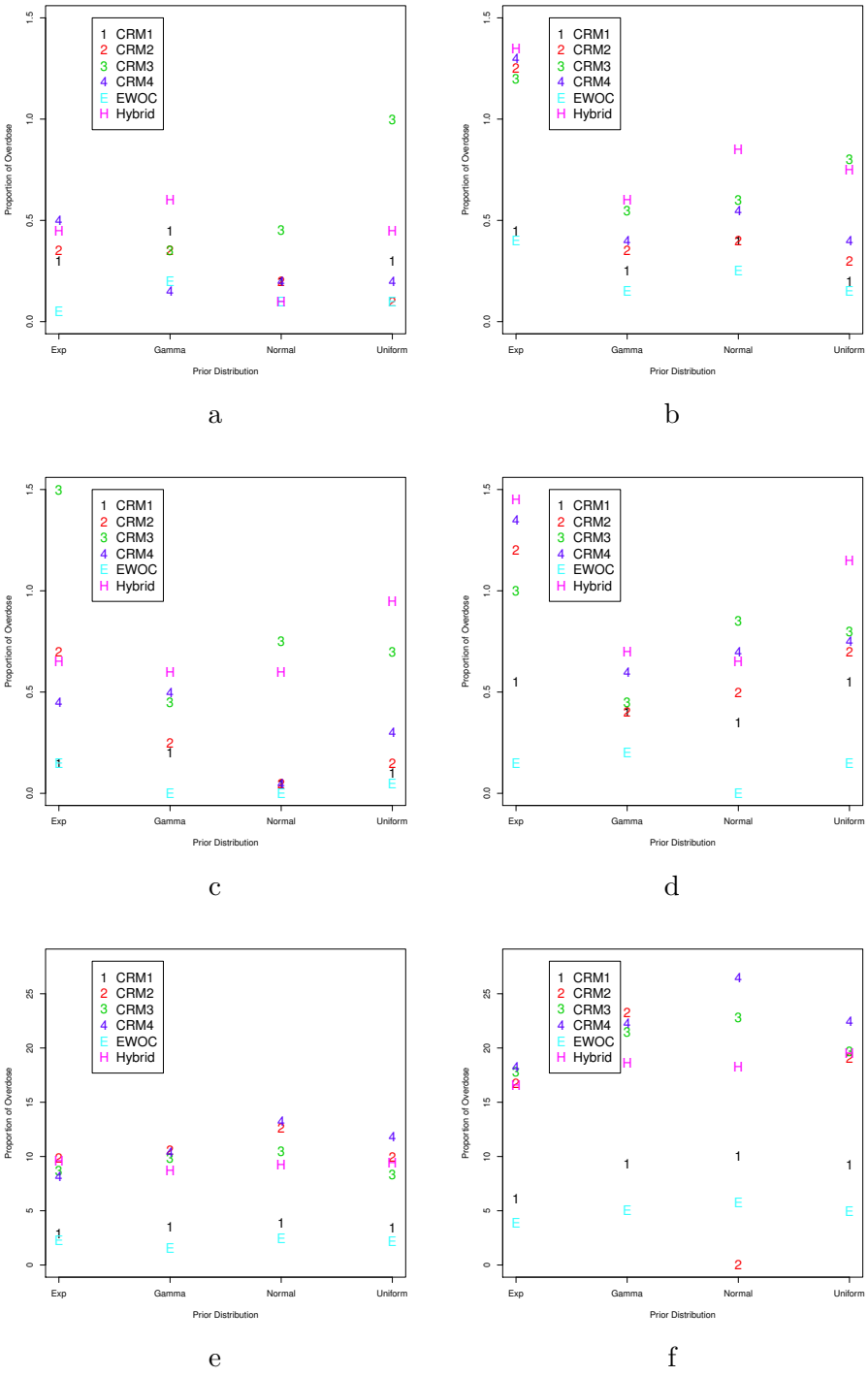


Figure 4.13: Overdose proportion with different prior distributions under different true Model. (a) true Model 1 (MTD_π), (b) true Model 1 (MTD_T), (c) true Model 2 (MTD_π), (d) true Model 2 (MTD_T), (e) true Model 3 (MTD_π), (f) true Model 3 (MTD_T).

4.1.5 Summary

Results in Model 1 and Model 2 are similar when the MTD level is medium (model 1) or high (model 2). For all the priors examined, all designs except EWOC have high proportions ($> 80\%$) of recommendation at the correct dose level. The hybrid design is comparable to CRM2, CRM3 and CRM4, and better than CRM1. The EWOC design has lower percentages of correct recommendation compared with other designs, and recommend a lower dose level (MTD-1) more often than the other designs. The performance using MTD_T is better than the corresponding ones using MTD_π. The percentages of recommendation at the overdose level(s) are very small ($< 1.8\%$) in all designs. The trend for the convergence rates is also similar. The convergence rates for all designs continued to increase as number of patients increased, but most of the gain had been obtained with 30 patients. The convergence rates for the hybrid design surpassed EWOC after 20 patients, and are similar to the CRM designs with 30 patients.

Under true model 3 with low MTD (MTD level $= 2$), all designs except EWOC with exponential prior have high proportions ($> 83\%$) of recommendation at the correct dose level using MTD_π. The EWOC design has lower percentages of correct recommendation with exponential prior, and recommends a lower dose level (MTD-1) more often than the other designs. When gamma prior or normal prior are used, the EWOC design has the comparable performance as the CRM1 design and is better than other designs; EWOC design is comparable to other designs and CRM1 design has the highest proportion of correct recommendation with uniform prior. When using MTD_T, the EWOC design has higher proportion of recommendation at the correct dose level than other designs for all priors considered. For both MTD_π and MTD_T, EWOC has the smallest proportions of recommendation at the overdose level for all the prior distributions examined. The hybrid design recommends fewer overdose levels than CRM2, CRM3, and CRM4 in most scenarios, and is comparable to CRM2, CRM3, or CRM4 in the rest scenarios. The EWOC design has higher proportion of correct recommendation using MTD_T than the corresponding ones using MTD_π in all priors examined except the gamma prior; whereas the performance for other designs using MTD_T is

worse than the corresponding ones using MTD_π for all priors examined except for CRM1 under exponential prior. The proportion of recommendation at the overdose level is higher using MTD_T than the corresponding ones using MTD_π. The performance based on convergence rates depends on the prior used. The EWOC design has the highest convergence rates or has the comparable high convergence rate with CRM1 design in all scenarios except the exponential and uniform prior using MTD_π, where the convergence rate for EWOC is lower than other designs. CRM1 has consistently high convergence rates in almost all scenarios.

4.2 Comparison to The Existing Models with High Baseline Hazard

The performance of the CATE model is compared to the TITE approach and the conventional approach under three true dose-response models with high baseline hazard (constant hazard $\lambda = 0.805$). Simulations were run for CRM, EWOC and hybrid designs using uniform prior distributions and varying the planned follow-up time (follow-up time window $\tau = 2$ vs. $\tau = 0.5$). CRM2 in Chu et al. (2009) are presented as the CRM model in the results.

The planned follow-up time of 2 is sufficient for the true dose-response models and the baseline hazard under consideration, where 86%, 89%, and 89% of all DLT on the respective MTD level would have been expected by time 2 under Model 1, Model 2, and Model 3.

The planned follow-up time of 0.5 is insufficient for the true dose-response models and the baseline hazard under consideration, where 43%, 43%, and 39% of all DLT on the respective MTD level would have been expected by time 0.5 under Model 1, Model 2, and Model 3. Under these scenarios, the true MTD should be defined using the MTD_π definition. When the planned follow up time is long enough, MTD_T converges to MTD_π and the MTD dose level under both definitions would be the same. When planned follow-up time is not long enough, the MTD_π definition is desired.

4.2.1 Model 1

The percentages of recommendation at each dose level under true model 1 are shown in Table 4.7. When the planned follow-up time is sufficient ($\tau = 2$), the proportion of recommendation at the correct dose level for the cure model using MTD_T (for all designs) is similar to that using the TITE approach and the conventional approach. The proportion of correct recommendation for the cure model using MTD_π is slightly lower than that using MTD_T for CRM or hybrid design, and about 10 percentage points lower for EWOC design. This might be explained by the fact that the probability of DLT under true model 1 at dose level 5 is closer to the target toxicity level θ at time 2 (correspond to the MTD_T definition) than at time 5 (correspond to the MTD_π definition since by time 5 the probability of DLT has reached a plateau) as depicted in Figure 3.2 (a). The EWOC design has lower percentages of correct recommendation than CRM or hybrid for all methods. The percentages of recommendation at the overdose level (level 6) are very small ($< 1.1\%$) in all scenarios. EWOC has lower percentages of recommendation at the overdose level than CRM or hybrid in most scenarios, and recommends a lower dose level ($MTD - 1$) more than CRM or hybrid. The convergence rates are shown in the upper panel of Figure 4.14. The convergence rates for all designs continued to increase as number of patients increased, but most of the gain had been obtained with 30 patients. The convergence rates for the CRM and hybrid designs under different methods are higher than the corresponding EWOC designs.

When the planned follow-up time is short ($\tau = 0.5$), the cure model using MTD_π maintains relatively high proportion of recommendation at the correct dose level, while the TITE and conventional approach overestimate the MTD and recommend the overdose level (dose level 6) most of the time. The cure model using MTD_T also recommends dose level 6 more frequently than the other doses, but to a lesser degree than TITE or conventional approach. As shown in Figure 3.2 (a), the probability of DLT still increases at time 0.5, and the planned follow up time of 0.5 is too short to observe most of the DLT for dose level 5 or 6 (only 43% of all DLT expected on MTD

level 5). The EWOC design has lower percentages of correct recommendation than CRM and hybrid for cure model using MTD_π, but has lower percentage of overdose, as it recommends a lower dose level ($MTD - 1$) more often than CRM and hybrid. Although TITE, conventional and cure model using MTD_T overestimate the MTD, the EWOC design tends to overestimate to a less degree than the corresponding CRM or hybrid design, and recommends the correct MTD level (level 5) more than CRM or hybrid. The convergence rates for cure model using MTD_π continued to increase as number of patients increased. All other methods did not converge as shown in the lower panel of Figure 4.14. When the planned follow-up time is insufficient and too short, selecting doses under cure model using the MTD_π definition is more appropriate than the MTD_T definition, and cure model using MTD_π is better than the TITE or conventional approach.

Table 4.7: Comparison Percentage of Recommended MTD by Dose Level Under True Model 1, Exponential Three-Parameter Model and $\lambda = 0.805$

Uniform prior $\tau = 2$							
Dose levels	1	2	3	4	5(MTD)*	6	Duration
CATE_π							
CRM	0.00	0.00	0.00	6.95	92.90	0.15	60.3
EWOC	0.00	0.00	0.00	23.70	76.25	0.05	60.5
Hybrid	0.00	0.00	0.00	7.25	92.60	0.15	60.6
CATE_T							
CRM	0.00	0.00	0.00	2.90	96.65	0.45	60.5
EWOC	0.00	0.00	0.00	13.45	86.40	0.15	60.4
Hybrid	0.00	0.00	0.00	3.25	95.70	1.05	60.3
TITE							
CRM	0.00	0.00	0.00	2.95	96.85	0.20	60.9
EWOC	0.00	0.00	0.00	13.90	85.85	0.25	60.8
Hybrid	0.00	0.00	0.00	3.80	95.15	1.05	60.8
Conventional							
CRM	0.00	0.00	0.00	2.65	96.85	0.50	111.1
EWOC	0.00	0.00	0.05	15.50	84.35	0.10	114.6
Hybrid	0.00	0.00	0.00	3.55	95.70	0.75	113.3
Uniform prior $\tau = 0.5$							
Dose levels	1	2	3	4	5(MTD)*	6	Duration
CATE_π							
CRM	0.00	0.00	0.00	8.10	78.95	12.95	59.4
EWOC	0.00	0.00	0.45	25.30	69.30	4.95	59.3
Hybrid	0.00	0.00	0.05	6.95	74.15	18.85	59.6
CATE_T							
CRM	0.00	0.00	0.00	0.05	28.75	71.20	59.3
EWOC	0.00	0.00	0.00	0.10	36.90	63.00	59.5
Hybrid	0.00	0.00	0.00	0.00	12.30	87.70	59.4
TITE							
CRM	0.00	0.00	0.00	0.00	10.70	89.30	59.3
EWOC	0.00	0.00	0.00	0.05	34.45	65.50	59.7
Hybrid	0.00	0.00	0.00	0.05	11.50	88.45	59.3
Conventional							
CRM	0.00	0.00	0.00	0.00	10.50	89.50	64.3
EWOC	0.00	0.00	0.00	0.10	35.90	64.00	64.8
Hybrid	0.00	0.00	0.00	0.00	11.55	88.45	64.5

*: The true MTD dose level was defined using MTD_π definition.

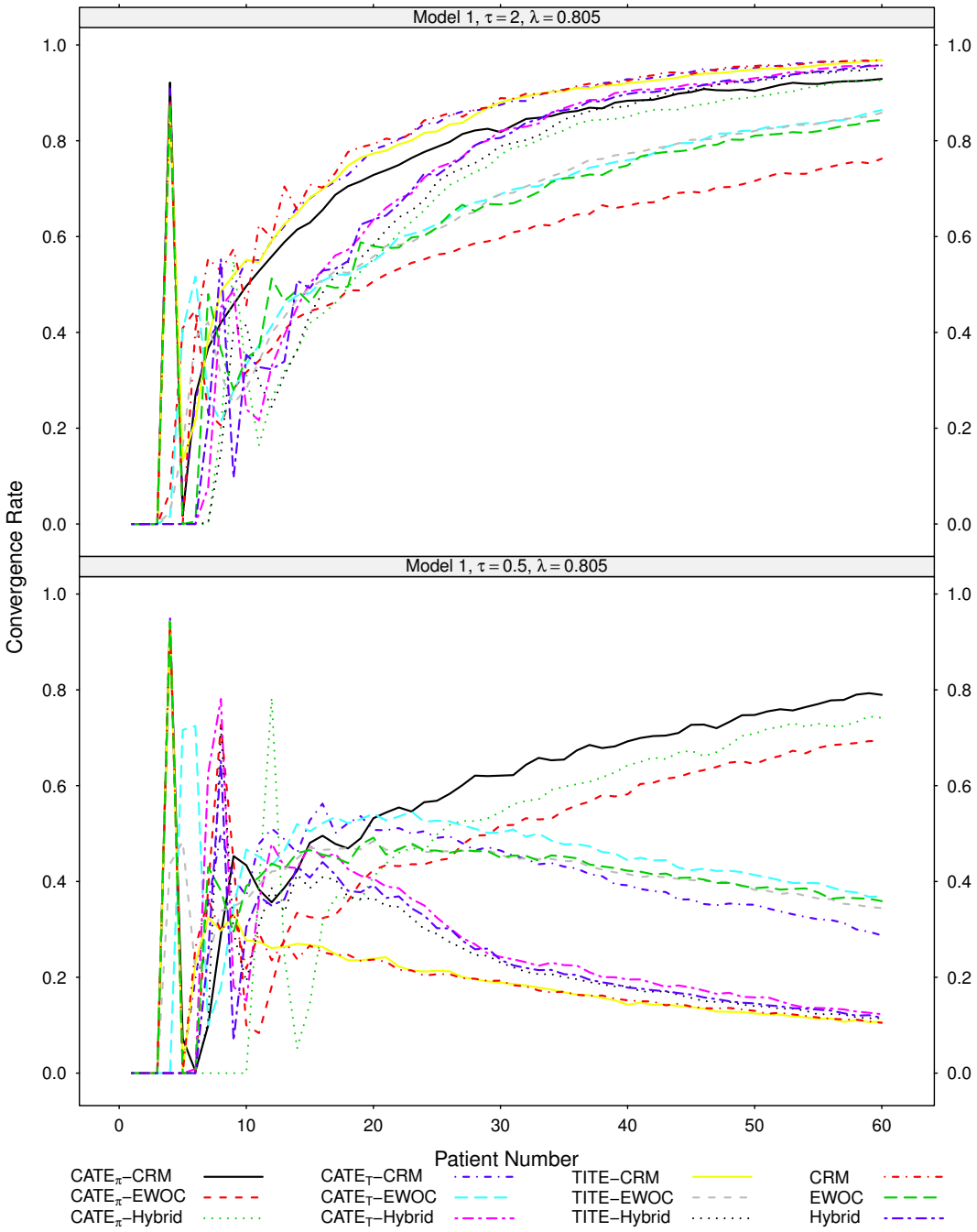


Figure 4.14: Convergence Rates with Different Methods Under True Model 1 When $\lambda = 0.805$

4.2.2 Model 2

The percentages of recommendation at each dose level under true model 2 are shown in Table 4.8. Similar results are observed under true model 2 as in true model 1. When the planned follow-up time τ equals to 2, the proportion of recommendation at the correct dose level for the cure model using MTD_T (for all designs) is similar to that using the TITE approach and the conventional approach. The proportion of correct recommendation for the cure model using MTD_π is slightly lower than that using MTD_T for CRM or hybrid design, and about 12 percentage points lower for EWOC design. As depicted in Figure 3.2 (b), the probability of DLT under true model 2 at dose level 4 is closer to the target toxicity level θ for the MTD_T definition (at time 2) than for the MTD_π definition (at time 5). The proportion of the correct recommendation is higher for CRM design than hybrid design than EWOC design for all approaches (cure model, TITE or conventional approach). The percentages of recommendation at the overdose level (level 5 or 6) are very small ($< 1.7\%$) in all scenarios. EWOC has lower percentages of recommendation at the overdose level than CRM or hybrid, and recommends a lower dose level ($MTD-1$) more than CRM or hybrid. The convergence rates are shown in the upper panel of Figure 4.15. The convergence rates for all designs continued to increase as number of patients increased, but most of the gain had been obtained with 30 patients. The convergence rates for the CRM and hybrid designs under different methods are higher than the corresponding EWOC designs.

When the planned follow-up time τ equals to 0.5, the cure model using MTD_π maintains relatively high proportion of recommendation at the correct dose level. The TITE and conventional approach overestimate the MTD and recommend the $MTD+1$ level (dose level 5) most of the time when the planned follow-up time is relatively short. The cure model using MTD_T also recommends dose level 5 more frequently than the other doses, but it recommends dose level 6 ($MTD + 2$) much lower than TITE or conventional approach. The probability of DLT in the true model continues to increase at time 0.5 as depicted in Figure 3.2 (b). The planned follow-up time of 0.5 is too short and only 43% of all DLT is expected on MTD level 4 by time 0.5. The EWOC

design has lower percentages of correct recommendation than CRM and hybrid for cure model using MTD_π, but has lower percentage of overdose, as it recommends a lower dose level ($MTD - 1$) more often than CRM or hybrid. Although TITE, conventional and cure model using MTD_T overestimate the MTD, the EWOC design has a higher percentages of recommendation at the correct dose level (level 4) than the corresponding CRM or hybrid design, and has a lower percentage of overdose recommendation. The convergence rates for cure model using MTD_π continued to increase as number of patients increased. All other methods did not converge as shown in the lower panel of Figure 4.15. When the planned follow-up time is insufficient and too short, selecting doses under cure model using the MTD_π definition is more appropriate and is better than the TITE or conventional approach.

Table 4.8: Comparison Percentage of Recommended MTD by Dose Level Under True Model 2, Exponential Three-Parameter Model and $\lambda = 0.805$

Uniform prior $\tau = 2$							
Dose levels	1	2	3	4(MTD)*	5	6	Duration
$CATE_\pi$							
CRM	0.00	0.00	11.45	88.30	0.25	0.00	60.6
EWOC	0.00	0.05	32.30	67.55	0.10	0.00	60.4
Hybrid	0.00	0.00	12.20	87.35	0.45	0.00	60.3
$CATE_T$							
CRM	0.00	0.00	5.00	94.40	0.60	0.00	60.6
EWOC	0.00	0.00	20.95	78.80	0.25	0.00	60.3
Hybrid	0.00	0.00	6.40	92.65	0.95	0.00	60.4
TITE							
CRM	0.00	0.00	4.75	94.15	1.10	0.00	60.6
EWOC	0.00	0.10	21.35	78.50	0.05	0.00	60.5
Hybrid	0.00	0.00	7.20	91.45	1.35	0.00	60.6
Conventional							
CRM	0.00	0.00	4.90	94.15	0.95	0.00	110.1
EWOC	0.00	0.00	22.85	76.95	0.20	0.00	113.6
Hybrid	0.00	0.00	6.50	91.85	1.65	0.00	112.7
Uniform prior $\tau = 0.5$							
Dose levels	1	2	3	4(MTD)*	5	6	Duration
$CATE_\pi$							
CRM	0.00	0.00	9.80	78.90	11.00	0.30	59.2
EWOC	0.05	1.25	30.20	65.85	2.60	0.05	59.2
Hybrid	0.00	0.10	9.65	75.55	14.25	0.45	59.4
$CATE_T$							
CRM	0.00	0.00	0.00	27.40	69.65	2.95	59.4
EWOC	0.00	0.00	0.30	40.45	56.95	2.30	59.6
Hybrid	0.00	0.00	0.00	14.20	73.45	12.35	59.1
TITE							
CRM	0.00	0.00	0.00	12.15	65.05	22.80	59.5
EWOC	0.00	0.00	0.55	39.15	54.65	5.65	59.5
Hybrid	0.00	0.00	0.05	18.70	64.05	17.20	59.5
Conventional							
CRM	0.00	0.00	0.00	14.25	62.75	23.00	64.2
EWOC	0.00	0.00	0.65	40.60	53.10	5.65	64.7
Hybrid	0.00	0.00	0.00	18.90	62.80	18.30	64.5

*: The true MTD dose level was defined using MTD_π definition.

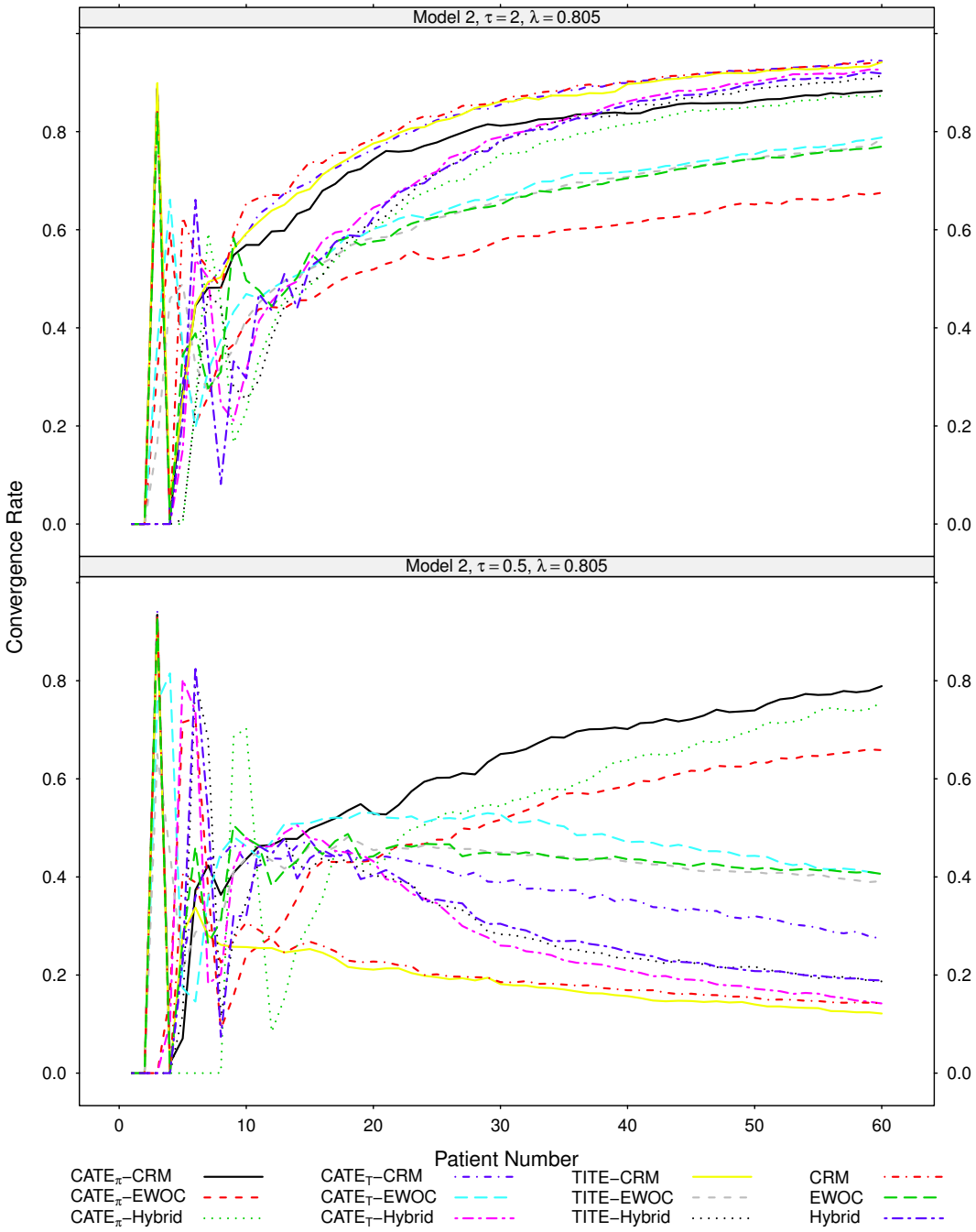


Figure 4.15: Convergence Rates with Different Methods Under True Model 2 When $\lambda = 0.805$

4.2.3 Model 3

Table 4.9 shows the percentages of recommendation at each dose level under true model 3. When τ equals to 2, the proportion of recommendation at the correct dose level for the cure model using MTD_T for EWOC design is similar to that using the TITE approach and the conventional approach, and the proportion for the cure model using MTD_π is lower than the TITE approach and the conventional approach. For CRM and hybrid design, the performance for the cure model using MTD_T is better than MTD_π, and both are better than the TITE approach or the conventional approach. The percentage of correct recommendation is higher for EWOC design than hybrid and CRM design for all approaches except for cure model using MTD_π, where EWOC design is comparable to CRM or hybrid (hybrid slightly higher followed by CRM). As shown in Figure 3.2 (c), the probability of DLT under true model 3 at dose level 2 is closer to the target toxicity level θ for the MTD_π definition (at time 5) than for the MTD_T definition (at time 2). EWOC has lower percentages of recommendation at the overdose level than CRM and hybrid, and recommends a lower dose level ($MTD-1$) more than CRM and hybrid. The convergence rates are shown in the upper panel of Figure 4.16. The convergence rates for all designs continued to increase as number of patients increased, but most of the gain had been obtained with 30 patients.

When the planned follow-up time is short ($\tau = 0.5$), the cure model using MTD_π has relatively high proportion of recommendation at the correct dose level. The cure model using MTD_T, TITE and conventional approach all overestimate the MTD, and recommend the $MTD+1$ level (dose level 3) most of the time. However, the cure model using MTD_T is less likely to recommend overdose levels above $MTD+1$ (dose level 4, 5 and 6) compared with TITE or conventional approach. As depicted in Figure 3.2 (c), the planned follow up time of 0.5 is too short and only 39% of all DLT is expected on MTD (level 2) by time 0.5. The EWOC design is comparable to CRM or hybrid in terms of correct recommendation for cure model using MTD_π, but has lower percentage of overdose, as it recommends a lower dose level ($MTD-1$) more often than CRM or hybrid. Although TITE, conventional and cure model using MTD_T overestimate the

MTD, the EWOC design has a higher percentages of recommendation at the correct dose level (level 2) than the corresponding CRM2 or hybrid design, and has a lower percentage of overdose recommendation. The convergence rates for cure model using MTD_π continued to increase as number of patients increased. All other methods did not converge as shown in the lower panel of Figure 4.16. Similar to what are observed under true model 1 or true model 2, when the planned follow-up time is too short to observe most of the DLT at the MTD level, cure model using the MTD_π is superior to other approaches.

Table 4.9: Comparison Percentage of Recommended MTD by Dose Level Under True Model 3, Exponential Three-Parameter Model and $\lambda = 0.805$

Uniform prior $\tau = 2$							
Dose levels	1	2(MTD)*	3	4	5	6	Duration
CATE_π							
CRM	1.70	87.05	11.25	0.00	0.00	0.00	60.8
EWOC	11.90	86.65	1.45	0.00	0.00	0.00	60.9
Hybrid	2.55	88.40	9.05	0.00	0.00	0.00	60.7
CATE_T							
CRM	0.35	78.70	20.95	0.00	0.00	0.00	60.8
EWOC	2.70	93.20	4.10	0.00	0.00	0.00	60.5
Hybrid	0.45	82.65	16.85	0.05	0.00	0.00	60.6
TITE							
CRM	0.25	76.10	23.65	0.00	0.00	0.00	60.5
EWOC	3.75	90.80	5.45	0.00	0.00	0.00	60.5
Hybrid	0.75	80.90	18.35	0.00	0.00	0.00	60.6
Conventional							
CRM	0.95	74.45	24.55	0.05	0.00	0.00	109.3
EWOC	3.50	90.30	5.75	0.00	0.00	0.00	113.7
Hybrid	0.70	80.20	19.10	0.00	0.00	0.00	112.5
Uniform prior $\tau = 0.5$							
Dose levels	1	2(MTD)*	3	4	5	6	Duration
CATE_π							
CRM	5.45	67.65	26.45	0.40	0.00	0.05	59.4
EWOC	24.65	66.50	8.80	0.05	0.00	0.00	59.4
Hybrid	7.30	64.65	27.00	1.05	0.00	0.00	59.5
CATE_T							
CRM	0.00	1.05	82.10	16.80	0.05	0.00	59.3
EWOC	0.00	7.60	84.85	7.55	0.00	0.00	59.3
Hybrid	0.00	1.00	75.45	23.20	0.35	0.00	59.4
TITE							
CRM	0.00	2.05	56.75	39.30	1.90	0.00	59.3
EWOC	0.00	13.80	71.75	14.40	0.05	0.00	59.3
Hybrid	0.00	4.80	68.20	26.35	0.65	0.00	59.5
Conventional							
CRM	0.00	1.95	57.80	38.55	1.70	0.00	64.3
EWOC	0.00	15.20	70.55	14.10	0.10	0.05	64.5
Hybrid	0.00	4.25	69.25	25.80	0.70	0.00	64.5

*: The true MTD dose level was defined using MTD_π definition.

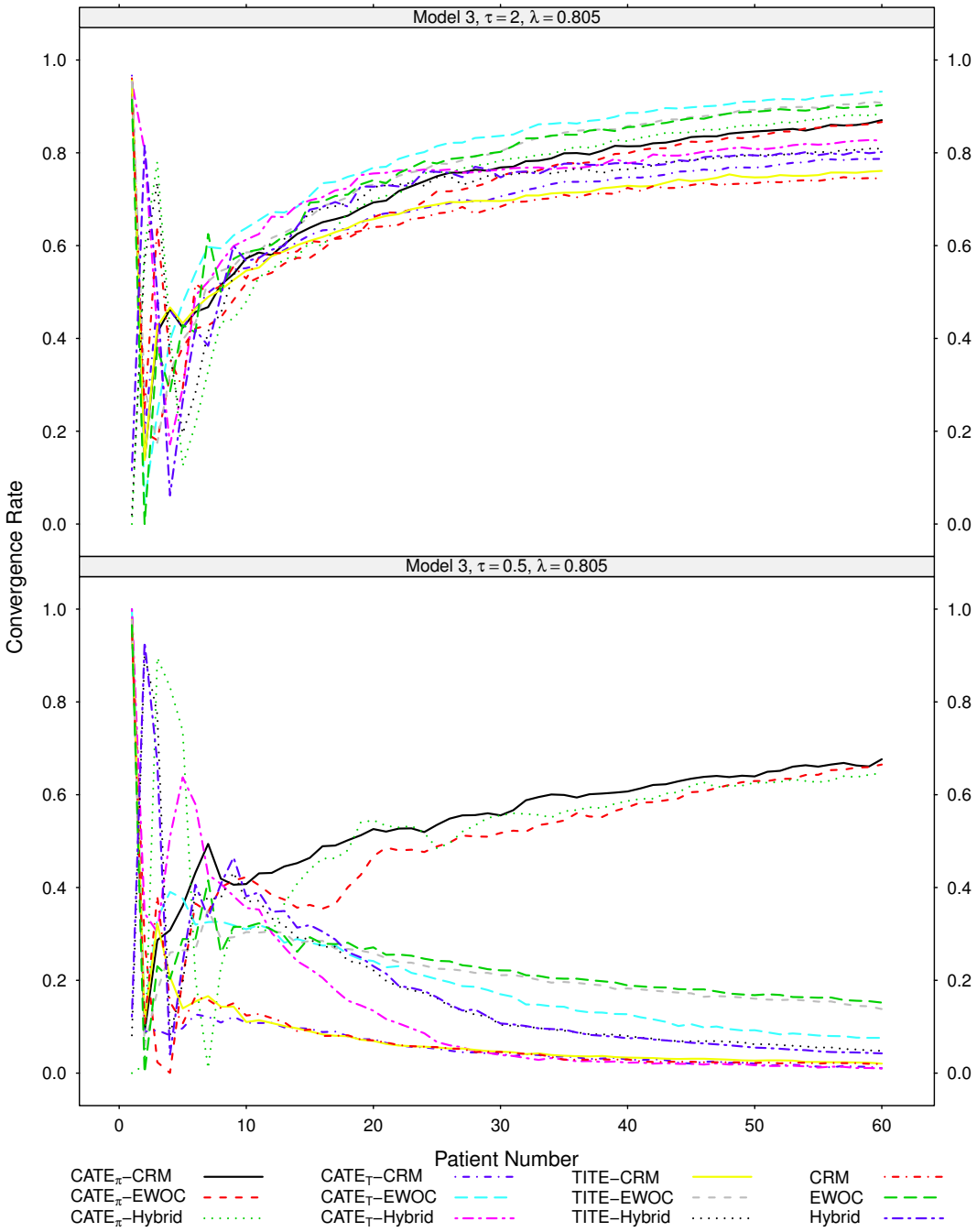


Figure 4.16: Convergence Rates with Different Methods Under True Model 3 When $\lambda = 0.805$

4.2.4 Summary

Results in Model 1 and Model 2 are similar when the MTD level is medium (model 1) or high (model 2). When the planned follow-up time is sufficient ($\tau = 2$, 86% to 89% of all DLT expected on MTD level by time 2), the proportion of recommendation at the correct dose level for the cure model using MTD_T is similar to that using the TITE approach and the conventional approach, and the percentage is lower for cure model using MTD_π compared with other approaches. The EWOC design has lower percentages of correct recommendation than CRM or hybrid for all approaches. The percentages of recommendation at the overdose level are very small ($< 1.7\%$) in all scenarios. EWOC is less likely to recommend the overdose level than CRM or hybrid, and recommends a lower dose level ($MTD - 1$) more than CRM or hybrid. The trend for the convergence rates is also similar. The convergence rates for all designs continued to increase as number of patients increased, but most of the gain had been obtained with 30 patients. The convergence rates for the CRM and hybrid designs are higher than the corresponding EWOC designs.

Under true model 3 with low MTD (MTD level = 2), when the planned follow-up time is sufficient (89% of all DLT expected on MTD level by time 2), the proportion of correct recommendation for EWOC design using cure model MTD_T is similar to that using the TITE approach and the conventional approach, and the proportion is lower for the cure model using MTD_π. For CRM and hybrid design, the performance for the cure model using MTD_T is better than MTD_π, and both are better than the TITE approach or the conventional approach. EWOC design is better than (in most cases) or at least comparable to CRM or hybrid (cure model using MTD_π) in terms of percentage of correct recommendation. EWOC has lower percentages of recommendation at the overdose level than CRM and hybrid, and recommends a lower dose level ($MTD - 1$) more than CRM and hybrid. Similar to model 1 and model 2, the convergence rates for all designs continued to increase as number of patients increased, but most of the gain had been obtained with 30 patients. The convergence rates for the EWOC designs are higher than the corresponding CRM and hybrid designs in most scenarios.

When the planned follow-up time is short ($\tau = 0.5$, only 39% to 43% of all DLT expected on MTD level by time 0.5), cure model using MTD_π is superior to other approaches for all true models considered. The cure model using MTD_π maintains relatively high proportion of recommendation at the correct dose level. The cure model using MTD_T, TITE and conventional approach all overestimate the MTD, and recommend the $MTD + 1$ level most of the time. However, the cure model using MTD_T is less likely to recommend overdose levels above $MTD + 1$ compared with TITE or conventional approach. The EWOC design has lower percentages of correct recommendation than CRM or hybrid for cure model using MTD_π under true model 1 or true model 2; while under true model 3, the EWOC design is comparable to CRM and hybrid. The EWOC design has a lower percentage of overdose recommendation than CRM and hybrid in all scenarios examined. The convergence rates for cure model using MTD_π continued to increase as number of patients increased. All other methods did not converge.

4.3 Comparison to The Existing Models with Low Baseline Hazard

The performance of the CATE model is compared to the TITE approach and the conventional approach under three true dose-response models with low baseline hazard (constant hazard $\lambda = 0.403$). Simulations were run for CRM, EWOC and hybrid designs using uniform prior distributions when the planned follow-up time $\tau = 2$. CRM2 in Chu et al. (2009) are presented as the CRM model in the results. This represents the case where the planned follow-up time is moderate relative to the risk of time to toxicity, where 67%, 67%, and 63% of all DLT on the respective MTD level would have been expected by time 2 under Model 1, Model 2, and Model 3, respectively.

4.3.1 Model 1

The percentages of recommendation at each dose level under true model 1 are shown in Table 4.10. The proportion of recommendation at the correct dose level for the cure model using MTD_T is better than that using the TITE approach or the conventional

approach for all designs. The cure model using MTD_π has higher correct recommendation percentage than that using MTD_T, TITE or conventional approach for CRM or hybrid design, but has lower correct recommendation percentage for EWOC design. Figure 3.3 (a) shows that the probability of DLT under true model 1 at dose level 5 is below the target toxicity level θ at time 2 (correspond to MTD_T) and slightly above the target toxicity level θ at time 5 (correspond to MTD_π). The percentage of correct recommendation for EWOC is lower than CRM or hybrid for cure model using MTD_π, while the percentage is higher than CRM and hybrid for all other approaches. EWOC has lower percentages of recommendation at the overdose level than CRM or hybrid, and recommends a lower dose level ($MTD - 1$) more than CRM or hybrid in all scenarios. The total trial duration using the cure model (both MTD_π and MTD_T) is much shorter compared with the conventional approach, and is similar to the TITE approach. The conventional approach is expected to have longer trial duration since a new patient has to wait until all current patients complete their follow-up before entering the trial, whereas the cure model and the TITE approach allow patients to enter the trial using the current information without suspending accrual. The convergence rates are shown in Figure 4.17. The convergence rates continued to increase as number of patients increased, and most of the gain had been obtained with 30 patients.

Table 4.10: Comparison Percentage of Recommended MTD by Dose Level Under True Model 1, Exponential Three-Parameter Model and $\lambda = 0.403$

Dose levels	1	2	3	4	5(MTD)	6	Duration
Uniform prior $\tau = 2$							
CATE_π							
CRM	0.00	0.00	0.00	7.70	90.20	2.10	60.5
EWOC	0.00	0.00	0.10	23.30	75.85	0.75	60.5
Hybrid	0.00	0.00	0.00	7.15	90.20	2.65	60.4
CATE_T							
CRM	0.00	0.00	0.00	0.45	86.75	12.80	60.5
EWOC	0.00	0.00	0.00	2.80	92.75	4.45	60.6
Hybrid	0.00	0.00	0.00	0.10	79.30	20.60	60.7
TITE							
CRM	0.00	0.00	0.00	0.15	76.75	23.10	60.7
EWOC	0.00	0.00	0.00	2.55	90.60	6.85	60.9
Hybrid	0.00	0.00	0.00	0.10	77.90	22.00	60.8
Conventional							
CRM	0.00	0.00	0.00	0.25	79.80	19.95	113.5
EWOC	0.00	0.00	0.00	3.05	91.40	5.55	116.6
Hybrid	0.00	0.00	0.00	0.40	77.90	21.70	115.2

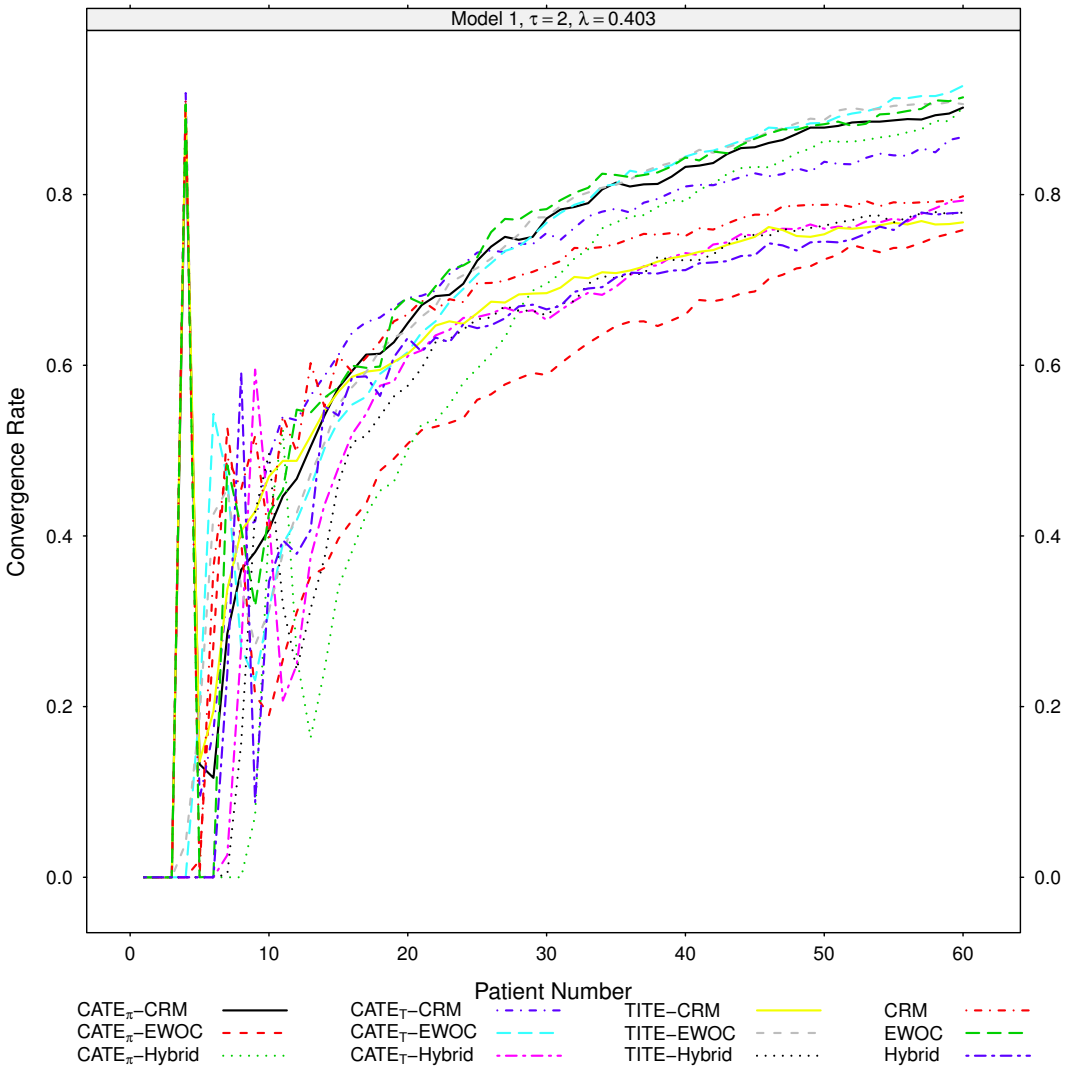


Figure 4.17: Convergence Rates with Different Methods Under True Model 1 When $\lambda = 0.403$

4.3.2 Model 2

Table 4.11 shows the percentages of recommendation at each dose level under true model 2. Results are similar to those under true model 1. The proportion of recommendation at the correct dose level for the cure model using MTD_T (for all designs) is higher than that using the TITE approach or the conventional approach. The proportion of correct recommendation for the cure model using MTD_π is higher for CRM or hybrid design, and lower for EWOC design than that using MTD_T, TITE or conventional approach. The percentage of correct recommendation for EWOC is lower than CRM or hybrid for cure model using MTD_π, while the percentage is higher than CRM or hybrid for all other approaches. EWOC has lower percentages of recommendation at the overdose level than CRM or hybrid, and recommends a lower dose level ($MTD - 1$) more than CRM or hybrid in all scenarios. The total trial duration using the cure model (both MTD_π and MTD_T) is much shorter compared with the conventional approach, and is similar to the TITE approach. The convergence rates continued to increase as number of patients increased (Figure 4.18).

Table 4.11: Comparison Percentage of Recommended MTD by Dose Level Under True Model 2, Exponential Three-Parameter Model and $\lambda = 0.403$

Dose levels	1	2	3	4(MTD)	5	6	Duration
Uniform prior $\tau = 2$							
$CATE_\pi$							
CRM	0.00	0.00	10.25	87.45	2.30	0.00	60.4
EWOC	0.00	0.50	31.40	67.60	0.50	0.00	60.8
Hybrid	0.00	0.00	10.25	86.50	3.05	0.20	60.6
$CATE_T$							
CRM	0.00	0.00	0.95	83.40	15.65	0.00	60.4
EWOC	0.00	0.05	5.90	89.40	4.60	0.05	60.4
Hybrid	0.00	0.00	0.85	81.90	17.25	0.00	60.5
TITE							
CRM	0.00	0.00	0.60	71.85	27.55	0.00	60.8
EWOC	0.00	0.00	5.70	86.90	7.35	0.05	60.9
Hybrid	0.00	0.00	1.15	75.65	23.15	0.05	60.9
Conventional							
CRM	0.00	0.00	0.80	71.75	27.35	0.10	112.0
EWOC	0.00	0.00	6.30	87.30	6.35	0.05	115.8
Hybrid	0.00	0.00	0.65	76.70	22.65	0.00	114.7

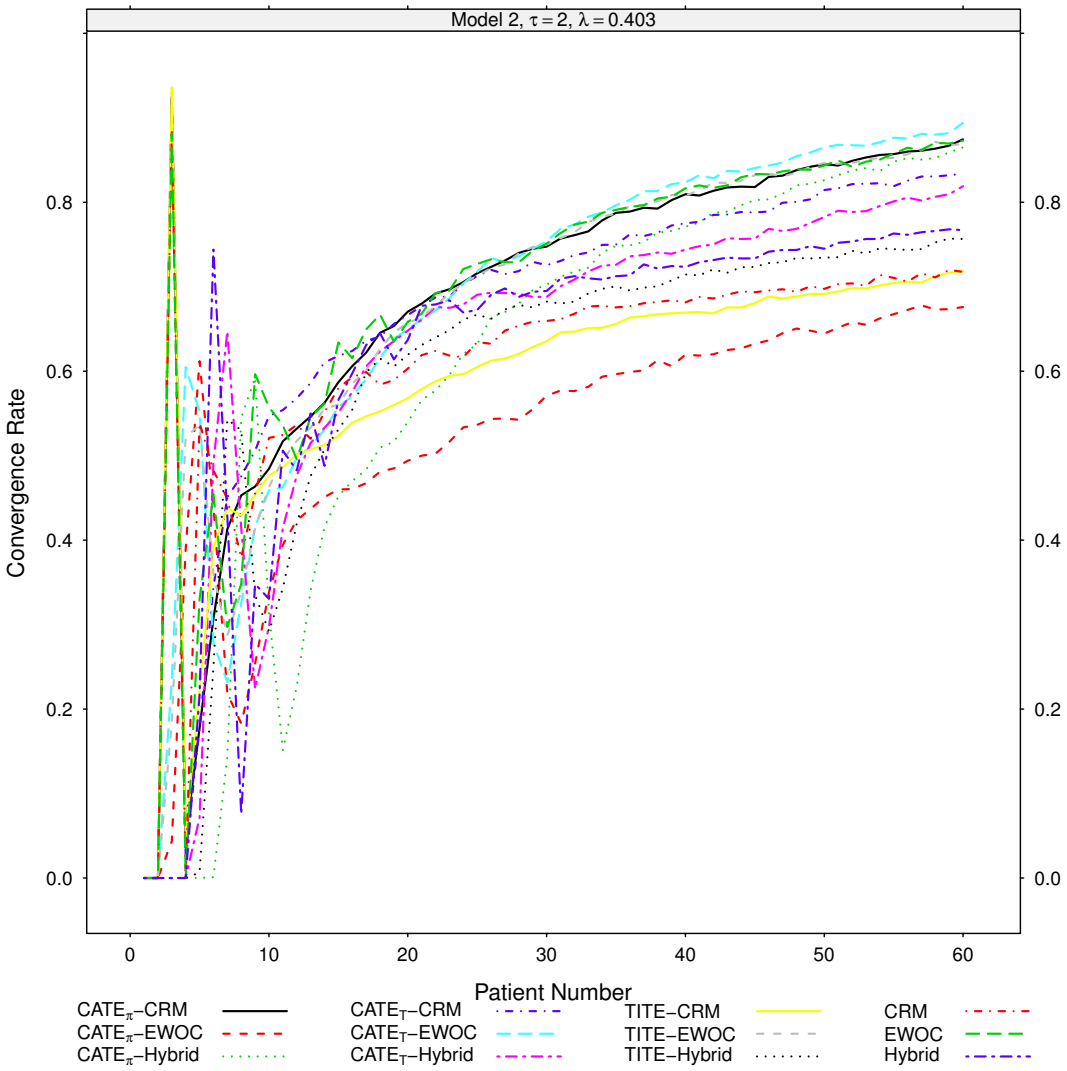


Figure 4.18: Convergence Rates with Different Methods Under True Model 2 When $\lambda = 0.403$

4.3.3 Model 3

The percentages of recommendation at each dose level under true model 3 are shown in table 4.12. The cure model using MTD_π has high proportion of recommendation at the correct dose level for both all designs and is superior to other approaches (cure model MTD_T, TITE or conventional approach). For EWOC design, the cure model using MTD_T recommends the correct dose level (dose level 2) most of the time; the

percentage is low than that using MTD_π, but higher than TITE or conventional approach. For CRM and hybrid design, TITE and conventional approach overestimate the MTD and recommends the $MTD + 1$ level most of the time. The cure model using MTD_T for CRM and hybrid design also recommends the $MTD + 1$ dose level more frequently than the other doses, but to a lesser degree than TITE or conventional approach. Figure 3.3 (c) shows that the probability of DLT under true model 3 at dose level 2 is closer to the target toxicity level θ at time 5 (correspond to MTD_π) than at time 2 (correspond to the MTD_T). The percentage of correct recommendation is higher for EWOC design than CRM or hybrid design for all approaches. EWOC has lower percentages of recommendation at the overdose level than CRM or hybrid. The total trial duration is similar for the cure model (both MTD_π and MTD_T) and the TITE approach, and is much shorter compared with the conventional approach. The convergence rates for cure model using MTD_π continued to increase as number of patients increased; the convergence rate for EWOC is lower at the beginning of the trial but accelerates to similar to CRM and hybrid with 40 patients (Figure 4.19). Convergence rates of EWOC for other approaches (cure model MTD_T, TITE or conventional) become stable with about 30 patients. CRM and hybrid for cure model using MTD_T, TITE or conventional approaches did not converge.

Table 4.12: Comparison Percentage of Recommended MTD by Dose Level Under True Model 3, Exponential Three-Parameter Model and $\lambda = 0.403$

Dose levels	1	2(MTD)	3	4	5	6	Duration
Uniform prior $\tau = 2$							
CATE_π							
CRM	3.00	78.85	18.15	0.00	0.00	0.00	60.5
EWOC	15.70	80.25	4.05	0.00	0.00	0.00	60.5
Hybrid	4.10	79.15	16.65	0.10	0.00	0.00	60.7
CATE_T							
CRM	0.00	36.80	63.05	0.15	0.00	0.00	60.7
EWOC	0.25	65.60	34.15	0.00	0.00	0.00	60.8
Hybrid	0.00	39.00	60.90	0.10	0.00	0.00	60.6
TITE							
CRM	0.00	27.90	71.40	0.70	0.00	0.00	60.7
EWOC	0.55	62.40	36.80	0.25	0.00	0.00	60.8
Hybrid	0.00	35.85	63.90	0.25	0.00	0.00	60.7
Conventional							
CRM	0.00	29.70	69.45	0.85	0.00	0.00	110.3
EWOC	0.65	62.05	37.25	0.05	0.00	0.00	114.9
Hybrid	0.05	37.10	62.45	0.40	0.00	0.00	113.4

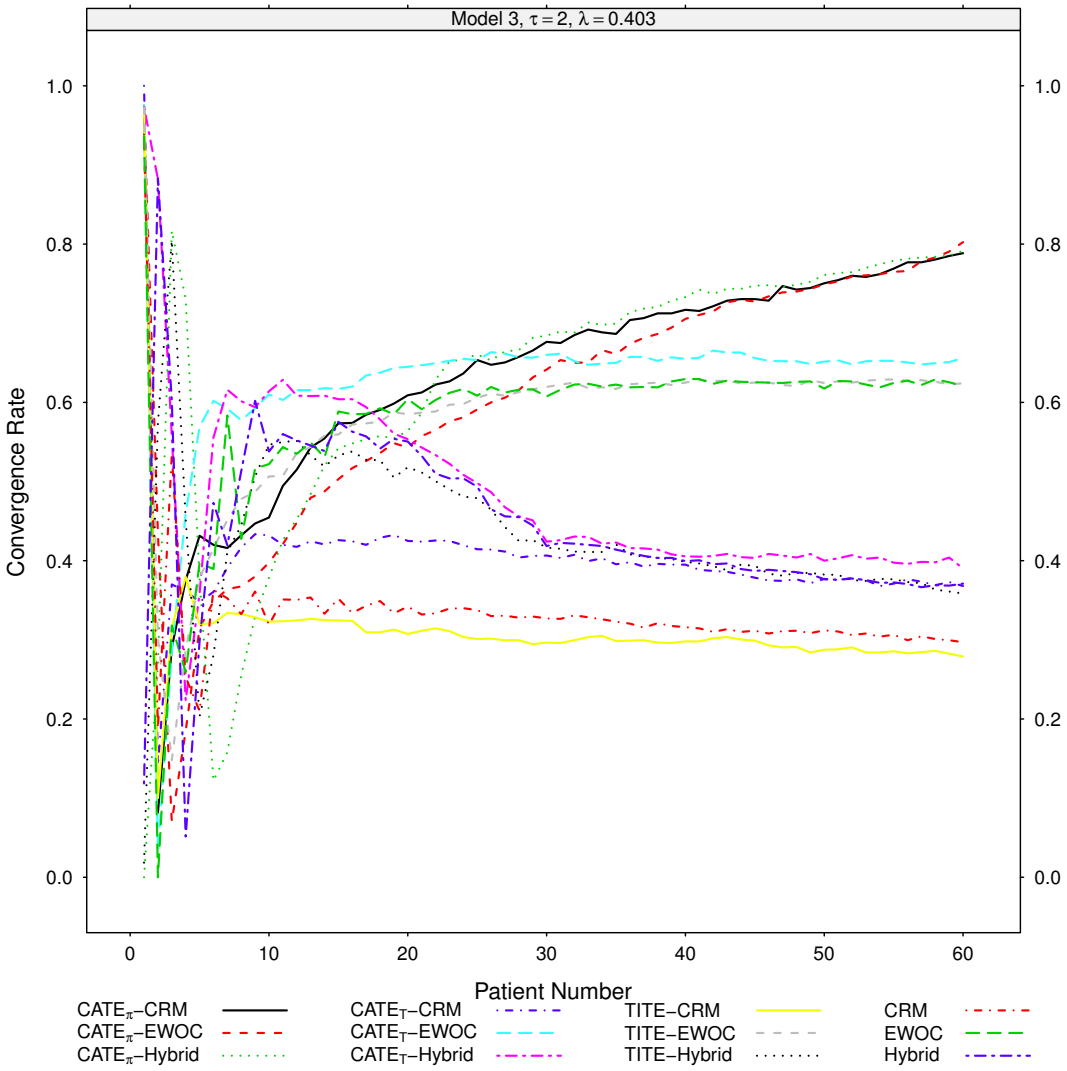


Figure 4.19: Convergence Rates with Different Methods Under True Model 3 When $\lambda = 0.403$

4.3.4 Summary

Results in Model 1 and Model 2 are similar when the MTD level is medium (model 1) or high (model 2). The planned follow-up time of 2 is moderate relative to the risk of time to toxicity, where 67% of all DLT on the MTD level would have been expected by time 2. The proportion of recommendation at the correct dose level for the cure model using MTD_T is better than that using the TITE approach or the conventional approach. The cure model using MTD_π has higher correct recommendation percentage than that using MTD_T, TITE or conventional approach for CRM and hybrid design, but has lower correct recommendation percentage for EWOC design. The percentage of correct recommendation for EWOC is lower than CRM or hybrid for cure model using MTD_π, while the percentage is higher than CRM or hybrid for all other approaches. EWOC has lower percentages of recommendation at the overdose level than CRM or hybrid, and recommends a lower dose level ($MTD - 1$) more than CRM or hybrid in all scenarios. The overdose percentage for cure model is lower than the corresponding ones for TITE or conventional approaches; and the overdose percentage of cure model using MTD_π is low than the corresponding MTD_T. The trend for the convergence rates is similar under true model 1 and model 2. The convergence rates continued to increase as number of patients increased, and most of the gain had been obtained with 30 patients.

Under true model 3 with low MTD (MTD level = 2), when the planned follow-up time is moderate (63% of all DLT expected on MTD level by time 2), the cure model using MTD_π is superior to other approaches (cure model MTD_T, TITE or conventional approach), and maintains high proportion of recommendation at the correct dose level. For EWOC design, the cure model using MTD_T recommends the correct dose level most of the time and is better than TITE or conventional approach. For CRM or hybrid design, the cure model using MTD_T, TITE and conventional approach all over-estimate the MTD and recommend the $MTD + 1$ level most of the time. However, the cure model using MTD_T is more likely to recommend the correct dose level and less likely to overdose compared with TITE or conventional approach. The percentage

of correct recommendation is higher for EWOC design than CRM or hybrid design for all approaches. EWOC has lower percentages of recommendation at the overdose level than CRM or hybrid. The convergence rates for cure model using MTD_π continued to increase as number of patients increased. Convergence rates of EWOC for other approaches become stable with about 30 patients. CRM and hybrid for cure model using MTD_T, TITE or conventional approaches did not converge.

For all true models, the total trial duration is similar for the cure model (both MTD_π and MTD_T) and the TITE approach, and is much shorter compared with the conventional approach.

4.4 Conclusion

It has been demonstrated by simulation that the proposed CATE desgin based on the proportional hazards model using exponential distribution with the known baseline hazard rate under different priors and true dose-toxicity models has generally high performance with high percentage of correct dose recommendation and low overdose proportion. When the true MTD is above the mid-level of the dose range considered (model 1 and model 2, the cases where the testing treatment is less toxic), the hybrid design and CRM designs have better convergence rates and recommend the target dose on MTD more often than EWOC. Designs using the MTD_T definition have better performance than the corresponding one using the MTD_π. The choice of prior has little effect on the performance of the designs. When the true MTD is below the mid-level of the dose range considered (model 3, the cases where the testing treatment is more toxic), EWOC design has better performance than other designs in most cases, but it somewhat depends on the prior distributions and the MTD definition. EWOC and the hybrid design generally provide a better safety protection in limiting higher dose for patients than the CRM designs do. Comparison between the two MTD definitions depends on the designs and the prior distributions. EWOC using the MTD_T definition generally have better performance than the corresponding one using the MTD_π, while the opposite is true for other designs.

The performance of the proposed CATE model is compared with TITE approach and the conventional approach under different planned follow-up time and true dose-toxicity models using the uniform prior distributions. The simulation study shows that the total trial duration is similar for the cure model (both MTD_π and MTD_T) and the TITE approach, and is much shorter compared with the conventional approach in all scenarios. The performance of the different approaches depends somewhat on the underlying dose-toxicity models, the planned follow-up time and the dose selecting designs (CRM and hybrid vs. EWOC).

When 86% to 89% of all DLT expected on MTD level by the planned follow-up time, the follow-up time is sufficient relative to the risk of time to toxicity. The performance for cure model using MTD_T is similar to that using the TITE approach or conventional approach, and is better than cure model using MTD_π when the true MTD is above the mid-level of the dose range considered. CRM and hybrid designs have better convergence rates and recommends the target dose on MTD more often than EWOC. When the true MTD is below the mid-level of the dose range considered (model 3), the performance for EWOC design using cure model MTD_T is similar to that using the TITE approach and the conventional approach, and is better than the cure model using MTD_π. For CRM and hybrid design, the performance for the cure model using MTD_T is better than MTD_π, and both are better than the TITE approach or the conventional approach. EWOC design has generally better performance than CRM or hybrid, and controls the overdose proportion better than CRM or hybrid.

When 63% to 67% of all DLT expected on MTD level by the planned follow-up time, the follow-up time is moderate relative to the risk of time to toxicity. When the true MTD is above the mid-level of the dose range considered, the performance for cure model MTD_T is better than TITE or conventional approach. For CRM and hybrid, the cure model using MTD_π has even better performance than that using MTD_T, TITE or conventional approach. CRM and hybrid designs have better convergence rates and recommend the target dose on MTD more often than EWOC in most cases. When the true MTD is below the mid-level of the dose range considered (model 3), the cure model using MTD_π is superior to other approaches and maintains high proportion

of recommendation at the correct dose level. For CRM and hybrid, TITE and the conventional approaches overestimate MTD and recommend $MTD + 1$ most of the time. The performance for cure model using MTD_T is better than TITE or conventional approach. The percentage of correct recommendation is higher for EWOC design than CRM or hybrid design for all approaches.

When the planned follow-up time is short (39% to 43% of all DLT expected on MTD level by the planned follow-up time), cure model using MTD_π is superior to other approaches for all true models considered. The cure model using MTD_π maintains relatively high proportion of recommendation at the correct dose level. TITE and conventional approach overestimate the MTD, and recommend the $MTD + 1$ level most of the time. Although the cure model using MTD_T also overestimates the MTD, it is less likely to recommend overdose levels above $MTD + 1$ compared with TITE or conventional approach. CRM and hybrid designs have better convergence rates and recommends the target dose on MTD more often than EWOC when the true MTD is above the mid-level of the dose range considered. When the true MTD is below the mid-level, EWOC design is comparable to CRM and hybrid. EWOC controls the overdose proportion better than CRM and hybrid for all scenarios.

In conclusion, the proposed cure model has generally high percentage of correct dose recommendation and low overdose proportion. It significantly reduces the overall trial duration compared with the conventional approach. When the planned follow-up time is moderate or too short relative to the risk of time to toxicity, the cure model (especially MTD_π for the short follow-up time) is superior to TITE or conventional approach, while TITE or conventional approach would overestimate MTD in certain situations.

Chapter 5

Performance of the CATE Design Based on the Proportional Hazards Model Using Exponential Distribution - Four-Parameter Model

The performance of the CATE Design based on the Proportional Hazards Model using exponential distribution where the baseline hazard rate λ is unknown is presented in this chapter. This is an exponential four-parameter models where $\alpha, \beta, \gamma, \lambda$ need to be estimated from data.

5.1 Performance of the CATE Design Using Exponential Four-Parameter Model

The performance of the proposed design is evaluated under three true dose-response models (low, middle, and high MTD) with high baseline hazard (hazard rate $\lambda = 0.805$) and with planned follow-up time $\tau = 2$. The planned follow-up time of 2 is sufficient for the true dose-response models and the baseline hazard under consideration, where 86%, 89%, and 89% of all DLT on the respective MTD level would have been expected by time 2 under Models 1, 2, and 3, respectively.

The percentages of recommendation at each dose level based on MTD_π and MTD_T for various designs, four different priors and three true models under the CATE design are summarized in Table 5.1, 5.2 and 5.3. CRM2 in Chu et al. (2009) are presented as the CRM model in the results.

Results presented here used four different priors for α, β, γ, and the informative gamma prior for λ, as described in Chapter 3.

5.1.1 Based on MTD_π

Table 5.1, 5.2 and 5.3 show that for all the models and the priors examined, both CRM and Hybrid designs under the CATE model have high proportions ($> 91\%, 81\%$ and 86% for Models 1, 2 and 3 respectively) of recommendation at the correct MTD dose level (Levels 5, 4 and 2 for Models 1, 2 and 3 respectively). The EWOC design has lower percentages of correct recommendation ($74\% - 83\%$ and $56\% - 70\%$ for Models 1 and 2 respectively) compared with other designs for Models 1 and 2 , and recommends a lower dose level more than the other designs. Compared with CRM and Hybrid, EWOC design for Model 3 has lower, comparable, a little higher, and a little higher percentages of correct recommendation for Exponential, Uniform, Gamma, and Normal priors, respectively. The hybrid design is comparable to CRM. The percentages of recommendation at the overdose level are very small ($\leq 0.8\%$) for all models and designs except for Model 3. For Model 3, the percentages of recommendation at the overdose levels are about 7 to 12% for CRM and Hybrid and is less than 3% for EWOC. The upper panels of Figure 5.1, 5.2, and 5.3 show the convergence rates based on MTD_π using the four different prior distributions under the three true models, respectively. The rates of convergences for CRM and Hybrid are higher than that for EWOC under Models 1 and 2 for all priors. The rates are comparable under Model 3 for all priors. The convergence rates for the hybrid design are similar to EWOC at the beginning of the trial, and accelerate to similar to the CRM designs after 30 patients. The convergence rates for all designs continue to increase as number of patients increases, but most of the gain has been obtained with 30 patients.

5.1.2 Based on MTD_T

Table 5.1, 5.2 and 5.3 also show that for Models 1 and 2 and all the priors examined, both CRM and Hybrid have very high proportions ($> 95\%$ and 90% for Models 1 and 2 respectively) of recommendation at the correct MTD dose level (Level 5 and 4 for Models 1 and 2 respectively). The performance using MTD_T is better than the corresponding ones using MTD_π for all designs. Similar to the cases using MTD_π, the

Table 5.1: Percentage of Recommended MTD by Dose Level Under True Model 1, Exponential Four-Parameter Model

				CATE_π		
Dose levels	1	2	3	4	5(MTD)	6
Exponential						
CRM	0.00	0.00	0.00	7.60	92.10	0.30
EWOC	0.00	0.00	0.05	26.25	73.50	0.20
Hybrid	0.00	0.00	0.00	7.65	91.65	0.70
Uniform						
CRM	0.00	0.00	0.00	7.20	92.65	0.15
EWOC	0.00	0.00	0.00	26.40	73.50	0.10
Hybrid	0.00	0.00	0.00	7.40	92.20	0.40
Gamma						
CRM	0.00	0.00	0.00	5.10	94.65	0.25
EWOC	0.00	0.00	0.00	16.75	83.20	0.05
Hybrid	0.00	0.00	0.00	4.55	94.85	0.60
Normal						
CRM	0.00	0.00	0.00	6.25	93.55	0.20
EWOC	0.00	0.00	0.00	21.15	78.70	0.15
Hybrid	0.00	0.00	0.00	5.75	93.65	0.60
				CATE_T		
Dose levels	1	2	3	4	5(MTD)	6
Exponential						
CRM	0.00	0.00	0.00	3.20	95.80	1.00
EWOC	0.00	0.00	0.05	14.10	85.50	0.35
Hybrid	0.00	0.00	0.00	2.85	96.10	1.05
Uniform						
CRM	0.00	0.00	0.00	3.35	96.15	0.50
EWOC	0.00	0.00	0.05	14.05	85.65	0.25
Hybrid	0.00	0.00	0.00	2.55	96.45	1.00
Gamma						
CRM	0.00	0.00	0.00	1.95	97.45	0.60
EWOC	0.00	0.00	0.00	9.00	90.75	0.25
Hybrid	0.00	0.00	0.00	1.55	97.25	1.20
Normal						
CRM	0.00	0.00	0.00	2.35	97.25	0.40
EWOC	0.00	0.00	0.00	8.95	90.75	0.30
Hybrid	0.00	0.00	0.00	2.40	97.15	0.45

Table 5.2: Percentage of Recommended MTD by Dose Level Under True Model 2, Exponential Four-Parameter Model

				CATE_π		
Dose levels	1	2	3	4(MTD)	5	6
Exponential						
CRM	0.00	0.00	16.80	82.95	0.25	0.00
EWOC	0.05	0.55	43.45	55.80	0.15	0.00
Hybrid	0.00	0.00	17.55	81.65	0.75	0.05
Uniform						
CRM	0.00	0.00	10.40	89.25	0.35	0.00
EWOC	0.00	0.20	35.45	64.25	0.10	0.00
Hybrid	0.00	0.00	12.95	86.55	0.45	0.05
Gamma						
CRM	0.00	0.00	9.80	90.05	0.15	0.00
EWOC	0.00	0.00	30.25	69.75	0.00	0.00
Hybrid	0.00	0.00	10.55	88.95	0.50	0.00
Normal						
CRM	0.00	0.00	9.30	90.50	0.20	0.00
EWOC	0.00	0.00	34.80	65.20	0.00	0.00
Hybrid	0.00	0.00	13.00	86.55	0.45	0.00
				CATE_T		
Dose levels	1	2	3	4(MTD)	5	6
Exponential						
CRM	0.00	0.00	9.20	90.10	0.70	0.00
EWOC	0.00	0.20	25.35	74.15	0.30	0.00
Hybrid	0.00	0.00	7.35	91.05	1.60	0.00
Uniform						
CRM	0.00	0.00	5.70	93.65	0.65	0.00
EWOC	0.00	0.05	23.50	76.40	0.05	0.00
Hybrid	0.00	0.00	6.70	92.45	0.85	0.00
Gamma						
CRM	0.00	0.00	4.50	95.00	0.50	0.00
EWOC	0.00	0.00	19.30	80.60	0.10	0.00
Hybrid	0.00	0.00	5.05	94.30	0.65	0.00
Normal						
CRM	0.00	0.00	5.80	93.75	0.45	0.00
EWOC	0.00	0.00	22.25	77.70	0.05	0.00
Hybrid	0.00	0.00	5.65	93.65	0.70	0.00

Table 5.3: Percentage of Recommended MTD by Dose Level Under True Model 3, Exponential Four-Parameter Model

			CATE_π			
Dose levels	1	2(MTD)	3	4	5	6
Exponential						
CRM	6.40	86.15	7.45	0.00	0.00	0.00
EWOC	23.25	75.75	1.00	0.00	0.00	0.00
Hybrid	5.70	86.55	7.75	0.00	0.00	0.00
Uniform						
CRM	1.75	87.55	10.70	0.00	0.00	0.00
EWOC	11.80	86.80	1.40	0.00	0.00	0.00
Hybrid	2.65	89.30	8.05	0.00	0.00	0.00
Gamma						
CRM	0.15	89.20	10.65	0.00	0.00	0.00
EWOC	2.75	95.55	1.70	0.00	0.00	0.00
Hybrid	0.55	91.75	7.70	0.00	0.00	0.00
Normal						
CRM	0.75	87.45	11.80	0.00	0.00	0.00
EWOC	6.90	90.20	2.90	0.00	0.00	0.00
Hybrid	1.10	89.20	9.70	0.00	0.00	0.00
			CATE_T			
Dose levels	1	2(MTD)	3	4	5	6
Exponential						
CRM	3.00	81.75	15.25	0.00	0.00	0.00
EWOC	11.80	84.55	3.65	0.00	0.00	0.00
Hybrid	1.90	81.80	16.20	0.10	0.00	0.00
Uniform						
CRM	0.25	80.90	18.85	0.00	0.00	0.00
EWOC	2.65	92.25	5.10	0.00	0.00	0.00
Hybrid	0.40	82.85	16.75	0.00	0.00	0.00
Gamma						
CRM	0.05	78.85	21.10	0.00	0.00	0.00
EWOC	0.60	94.85	4.55	0.00	0.00	0.00
Hybrid	0.10	81.90	18.00	0.00	0.00	0.00
Normal						
CRM	0.00	76.45	23.55	0.00	0.00	0.00
EWOC	0.70	91.80	7.50	0.00	0.00	0.00
Hybrid	0.15	81.00	18.85	0.00	0.00	0.00

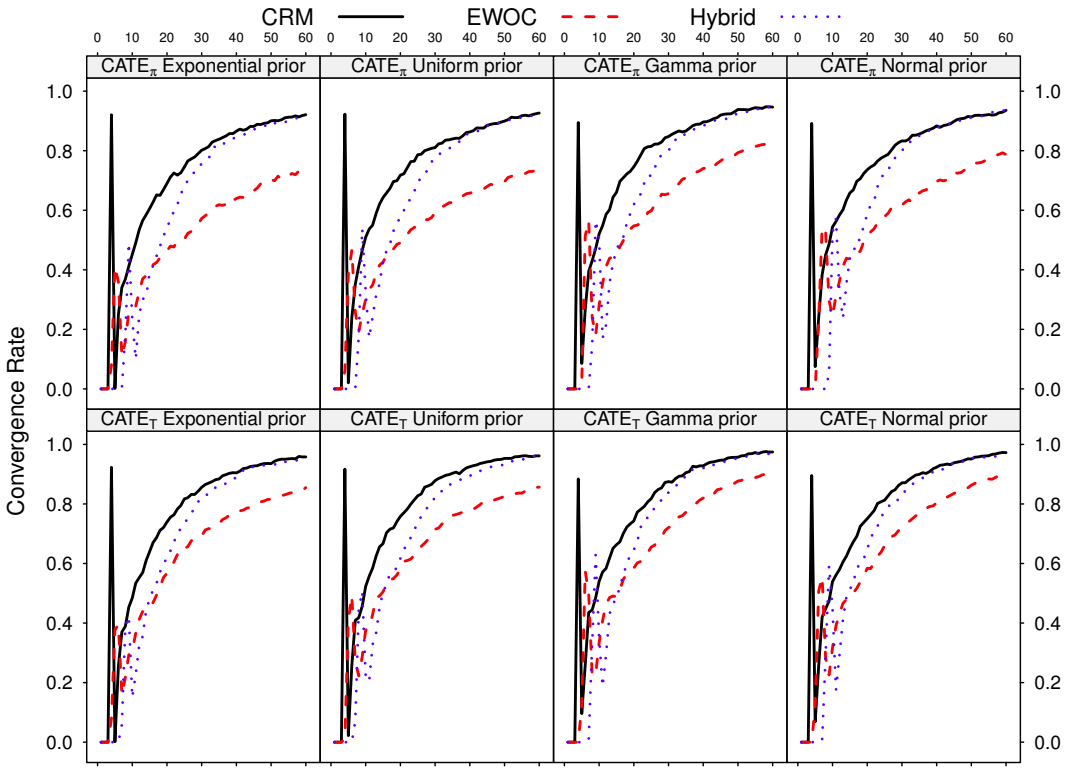


Figure 5.1: Convergence Rates Under True Model 1

EWOC design has slightly lower percentages of correct recommendation ($85\% - 91\%$ and $74\% - 81\%$ for Models 1 and 2 respectively) compared with CRM and Hybrid, and recommend a lower dose level more often than the other designs. The hybrid design is comparable to CRM. The percentages of recommendation at the overdose level are very small ($\leq 1.6\%$) in all designs. For Model 3, EWOC design has higher proportion of correct recommendation (true MTD of level 2) than CRM or Hybrid. CRM and Hybrid tends to have a higher recommendation to the $MTD + 1$ level (Level 3) compared to EWOC. One of the reasons for EWOC performed better in this scenario is that the true MTD occurs between dose levels 2 and 3 from the calculation based on MTD_T under Model 3. The lower panels of Figure 5.1, 5.2, and 5.3 show the convergence rates based on MTD_T using the four different prior distributions under the three true models, respectively. The trend for the convergence rates is similar to those for MTD_π and most of gain has been obtained with 30 patients under all the true models and the priors .

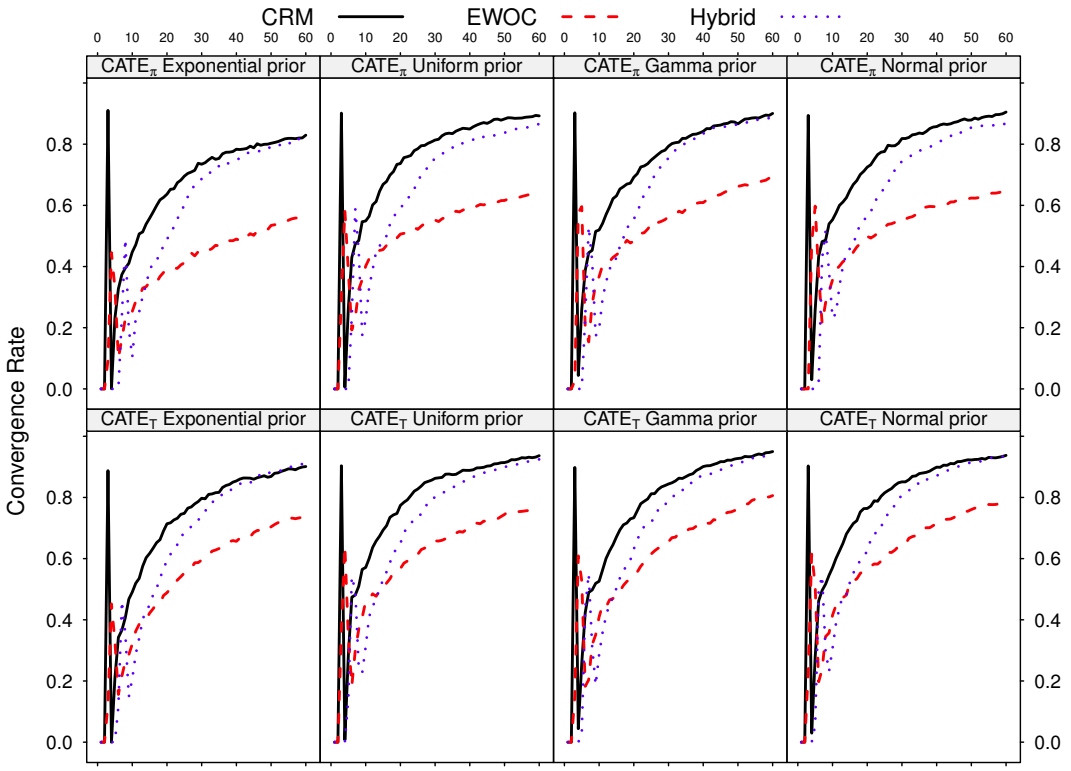


Figure 5.2: Convergence Rates Under True Model 2

5.1.3 Comparison Between MTD_π and MTD_T

The performance using MTD_T is a little better than the corresponding ones using MTD_π for Models 1 and 2 and all the priors examined, although CRM and Hybrid have high proportions of recommendation at the correct MTD dose level than EWOC based on either MTD_π or MTD_T. The EWOC design has slightly lower percentages of correct recommendation compared with other designs, and recommend a lower dose level more often than the other designs. For Model 3, it seems that using MTD_π performs better than using MTD_T for both CRM and Hybrid. EWOC using MTD_T performs better than that using MTD_π because the true MTD lies between dose level 2 and 3 when MTD_T is used.

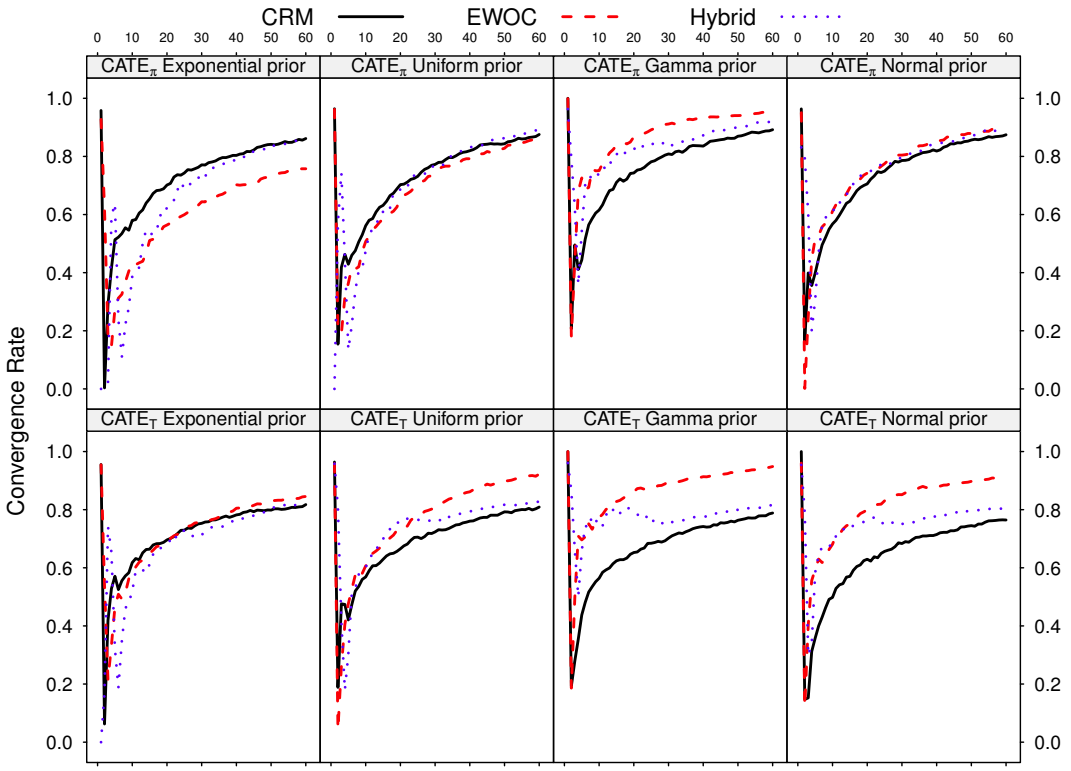


Figure 5.3: Convergence Rates Under True Model 3

5.1.4 Overdose Comparisons Under Different True Models and Different Prior Distributions

The percentage of overdose at recommended dose level for each design under different true models and different prior distributions is depicted in Figure 5.4. EWOC design has the lowest percentages in all scenarios. The percentages of overdose for all designs are very low under true model 1 or true model 2 ($\leq 1.6\%$ for all priors). Under true model 3, the hybrid design recommends fewer overdose levels than CRM does for all priors except the exponential prior where the overdose proportions for hybrid and CRM are very similar.

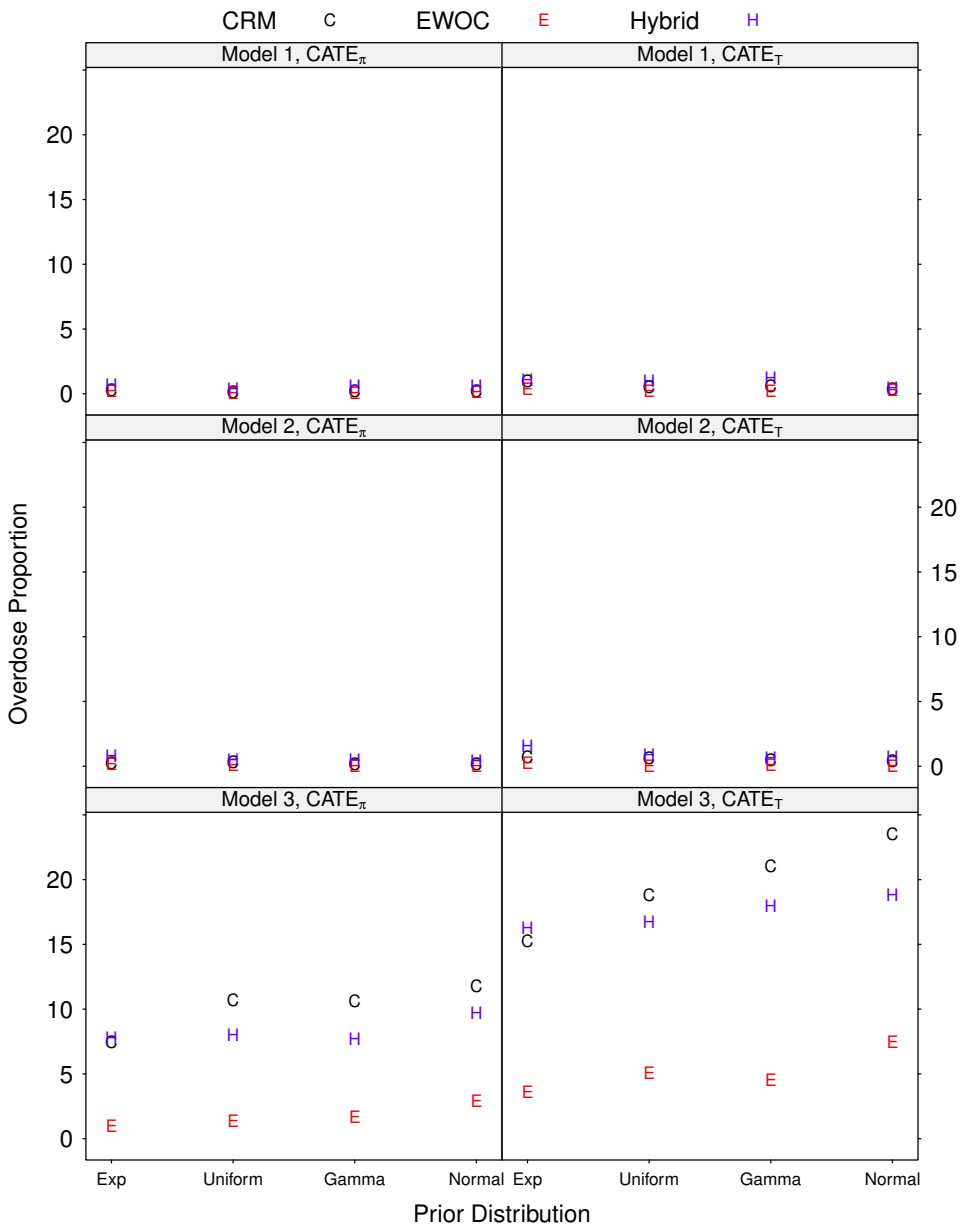


Figure 5.4: Overdose Proportion with Different Prior Distributions Under Different True Models

5.2 Performance of the CATE Design by Comparing to Other Existing Designs

The performance of the CATE design is compared to the TITE approach and the conventional approach under three true dose-response models with high baseline hazard (hazard rate $\lambda = 0.805$). Simulations were run for CRM, EWOC and hybrid designs using uniform prior distributions for α, β, γ, and the informative gamma prior for λ, and varying the planned follow-up time (follow-up time window $\tau = 2$ vs. $\tau = 0.5$).

The planned follow up time of 2 is sufficient for the true dose-response models and the baseline hazard under consideration, where 86%, 89%, and 89% of all DLT on the respective MTD level would have been expected by time 2 under Model 1, Model 2, and Model 3.

The planned follow up time of 0.5 is insufficient for the true dose-response models and the baseline hazard under consideration, where 43%, 43%, and 39% of all DLT on the respective MTD level would have been expected by time 0.5 under Model 1, Model 2, and Model 3. Under these scenarios, the true MTD should be defined using the MTD_π definition. The percentages of recommendation at each dose level under the three true models are shown in Table 5.4, 5.5, and 5.6, respectively.

5.2.1 When the Planned Follow-up Time is Sufficient ($\tau = 2$)

For Model 1, as shown in Table 5.4, when the planned follow-up time is sufficient ($\tau = 2$), the proportion of recommendation at the correct dose level for the CATE design using MTD_T (for all designs) is similar to that using the TITE approach and the conventional approach. The proportion of correct recommendation for the CATE design using MTD_π is slightly lower than that using MTD_T for CRM or hybrid design, and about 12 percentage points lower for EWOC design. This might be explained by the fact that the probability of DLT under true model 1 at dose level 5 is closer to the target toxicity level θ at time 2 (corresponding to the MTD_T definition) than at time 5 (corresponding to the MTD_π definition since by time 5 the probability of DLT has reached a plateau) as depicted in Figure 3.2a. The EWOC design has lower percentages

of correct recommendation than CRM or hybrid for all methods. The percentages of recommendation at the overdose level (level 6) are very small ($< 1.1\%$) in all scenarios.

For Model 2, as shown in Table 5.5, when the planned follow-up time τ equals to 2, the proportion of recommendation at the correct dose level for the CATE design using MTD_T (for all designs) is similar to that using the TITE approach and the conventional approach. The proportion of correct recommendation for the CATE design using MTD_π is slightly lower than that using MTD_T for CRM or hybrid design, and about 12 percentage points lower for EWOC design. As depicted in Figure 3.2b, the probability of DLT under true model 2 at dose level 4 is closer to the target toxicity level θ for the MTD_T definition (at time 2) than for the MTD_π definition. The proportion of the correct recommendation is higher for CRM or hybrid design than EWOC design for all approaches (CATE, TITE or conventional approaches). The percentages of recommendation at the overdose level (level 5 or 6) are very small ($< 1.7\%$) in all scenarios. EWOC has lower percentages of overdose recommendation than CRM or hybrid, and recommends a lower dose level ($MTD - 1$) more than CRM.

For Model 3, as shown in Table 5.6, when τ equals to 2, the proportion of recommendation at the correct dose level for the CATE model using MTD_T for EWOC design is slightly higher than the TITE approach and the conventional approach, and the proportion for the CATE model using MTD_π is lower than the TITE approach and the conventional approach. For CRM or hybrid design, the performance for the CATE model (both MTD_T and MTD_π) is better than the TITE approach or the conventional approach. The proportion of correct recommendation for the CATE model using MTD_π is higher than that using MTD_T for CRM or hybrid design, and lower for EWOC design. The percentage of correct recommendation is higher for EWOC design than CRM or hybrid design for all approaches except for CATE model using MTD_π, where EWOC design is comparable to CRM or hybrid. As shown in Figure 3.2c, the probability of DLT under true model 3 at dose level 2 is closer to the target toxicity level θ for the MTD_π definition than for the MTD_T definition (at time 2). EWOC has lower percentages of recommendation at the overdose level than CRM or hybrid, and recommends a lower dose level ($MTD - 1$) more than CRM or hybrid.

The convergence rates are shown in the upper panels of Figure 5.5. For all the models, the convergence rates for all designs continue to increase as number of patients increases, but most of the gain has been obtained with 30 patients. The convergence rates for the CRM or hybrid designs under different methods are higher than the corresponding EWOC designs for Model 1 or Model 2.

Table 5.4: Comparison Percentage of Recommended MTD by Dose Level Under True Model 1, Exponential Four-Parameter Model with Informative Prior on λ and $\lambda = 0.805$

				$\tau = 2$			
Dose levels	1	2	3	4	5(MTD)	6	Duration
CATE_π							
CRM	0.00	0.00	0.00	7.20	92.65	0.15	60.8
EWOC	0.00	0.00	0.00	26.40	73.50	0.10	60.6
Hybrid	0.00	0.00	0.00	7.40	92.20	0.40	60.6
CATE_T							
CRM	0.00	0.00	0.00	3.35	96.15	0.50	60.7
EWOC	0.00	0.00	0.05	14.05	85.65	0.25	60.8
Hybrid	0.00	0.00	0.00	2.55	96.45	1.00	60.5
TITE							
CRM	0.00	0.00	0.00	2.95	96.85	0.20	60.9
EWOC	0.00	0.00	0.00	13.90	85.85	0.25	60.8
Hybrid	0.00	0.00	0.00	3.80	95.15	1.05	60.8
Conventional							
CRM	0.00	0.00	0.00	2.65	96.85	0.50	111.1
EWOC	0.00	0.00	0.05	15.50	84.35	0.10	114.6
Hybrid	0.00	0.00	0.00	3.55	95.70	0.75	113.3
				$\tau = 0.5$			
Dose levels	1	2	3	4	5(MTD)	6	Duration
CATE_π							
CRM	0.00	0.05	0.05	9.30	78.30	12.30	59.5
EWOC	0.00	0.00	0.45	28.90	66.40	4.25	59.5
Hybrid	0.00	0.00	0.00	7.85	74.00	18.15	59.7
CATE_T							
CRM	0.00	0.00	0.00	0.00	32.15	67.85	59.6
EWOC	0.00	0.00	0.00	0.10	34.65	65.25	59.5
Hybrid	0.00	0.00	0.00	0.00	12.00	88.00	59.6
TITE							
CRM	0.00	0.00	0.00	0.00	10.70	89.30	59.3
EWOC	0.00	0.00	0.00	0.05	34.45	65.50	59.7
Hybrid	0.00	0.00	0.00	0.05	11.50	88.45	59.3
Conventional							
CRM	0.00	0.00	0.00	0.00	10.50	89.50	64.3
EWOC	0.00	0.00	0.00	0.10	35.90	64.00	64.8
Hybrid	0.00	0.00	0.00	0.00	11.55	88.45	64.5

Table 5.5: Comparison Percentage of Recommended MTD by Dose Level Under True Model 2, Exponential Four-Parameter Model with Informative Prior on λ and $\lambda = 0.805$

				$\tau = 2$			
Dose levels	1	2	3	4(MTD)	5	6	Duration
CATE_π							
CRM	0.00	0.00	10.40	89.25	0.35	0.00	60.4
EWOC	0.00	0.20	35.45	64.25	0.10	0.00	60.6
Hybrid	0.00	0.00	12.95	86.55	0.45	0.05	60.9
CATE_T							
CRM	0.00	0.00	5.70	93.65	0.65	0.00	60.6
EWOC	0.00	0.05	23.50	76.40	0.05	0.00	60.8
Hybrid	0.00	0.00	6.70	92.45	0.85	0.00	60.8
TITE							
CRM	0.00	0.00	4.75	94.15	1.10	0.00	60.6
EWOC	0.00	0.10	21.35	78.50	0.05	0.00	60.5
Hybrid	0.00	0.00	7.20	91.45	1.35	0.00	60.6
Conventional							
CRM	0.00	0.00	4.90	94.15	0.95	0.00	110.1
EWOC	0.00	0.00	22.85	76.95	0.20	0.00	113.6
Hybrid	0.00	0.00	6.50	91.85	1.65	0.00	112.7
				$\tau = 0.5$			
Dose levels	1	2	3	4(MTD)	5	6	Duration
CATE_π							
CRM	0.00	0.20	12.20	78.15	9.25	0.20	59.3
EWOC	0.00	1.65	31.20	64.30	2.80	0.05	59.7
Hybrid	0.00	0.00	11.50	72.70	15.05	0.75	59.6
CATE_T							
CRM	0.00	0.00	0.00	29.90	66.90	3.20	59.4
EWOC	0.00	0.00	0.05	41.80	55.90	2.25	59.5
Hybrid	0.00	0.00	0.00	13.60	73.20	13.20	59.8
TITE							
CRM	0.00	0.00	0.00	12.15	65.05	22.80	59.5
EWOC	0.00	0.00	0.55	39.15	54.65	5.65	59.5
Hybrid	0.00	0.00	0.05	18.70	64.05	17.20	59.5
Conventional							
CRM	0.00	0.00	0.00	14.25	62.75	23.00	64.2
EWOC	0.00	0.00	0.65	40.60	53.10	5.65	64.7
Hybrid	0.00	0.00	0.00	18.90	62.80	18.30	64.5

Table 5.6: Comparison Percentage of Recommended MTD by Dose Level Under True Model 3, Exponential Four-Parameter Model with Informative Prior on λ and $\lambda = 0.805$

			$\tau = 2$				
Dose levels	1	2(MTD)	3	4	5	6	Duration
CATE_π							
CRM	1.75	87.55	10.70	0.00	0.00	0.00	60.9
EWOC	11.80	86.80	1.40	0.00	0.00	0.00	60.7
Hybrid	2.65	89.30	8.05	0.00	0.00	0.00	60.7
CATE_T							
CRM	0.25	80.90	18.85	0.00	0.00	0.00	60.4
EWOC	2.65	92.25	5.10	0.00	0.00	0.00	60.7
Hybrid	0.40	82.85	16.75	0.00	0.00	0.00	60.8
TITE							
CRM	0.25	76.10	23.65	0.00	0.00	0.00	60.5
EWOC	3.75	90.80	5.45	0.00	0.00	0.00	60.5
Hybrid	0.75	80.90	18.35	0.00	0.00	0.00	60.6
Conventional							
CRM	0.95	74.45	24.55	0.05	0.00	0.00	109.3
EWOC	3.50	90.30	5.75	0.00	0.00	0.00	113.7
Hybrid	0.70	80.20	19.10	0.00	0.00	0.00	112.5
			$\tau = 0.5$				
Dose levels	1	2(MTD)	3	4	5	6	Duration
CATE_π							
CRM	5.35	67.35	26.60	0.70	0.00	0.00	59.5
EWOC	26.50	66.50	6.90	0.10	0.00	0.00	59.4
Hybrid	6.90	65.20	27.40	0.50	0.00	0.00	59.7
CATE_T							
CRM	0.00	2.20	83.90	13.85	0.05	0.00	59.5
EWOC	0.00	7.00	84.95	8.00	0.05	0.00	59.4
Hybrid	0.00	0.65	74.60	24.55	0.20	0.00	59.6
TITE							
CRM	0.00	2.05	56.75	39.30	1.90	0.00	59.3
EWOC	0.00	13.80	71.75	14.40	0.05	0.00	59.3
Hybrid	0.00	4.80	68.20	26.35	0.65	0.00	59.5
Conventional							
CRM	0.00	1.95	57.80	38.55	1.70	0.00	64.3
EWOC	0.00	15.20	70.55	14.10	0.10	0.05	64.5
Hybrid	0.00	4.25	69.25	25.80	0.70	0.00	64.5

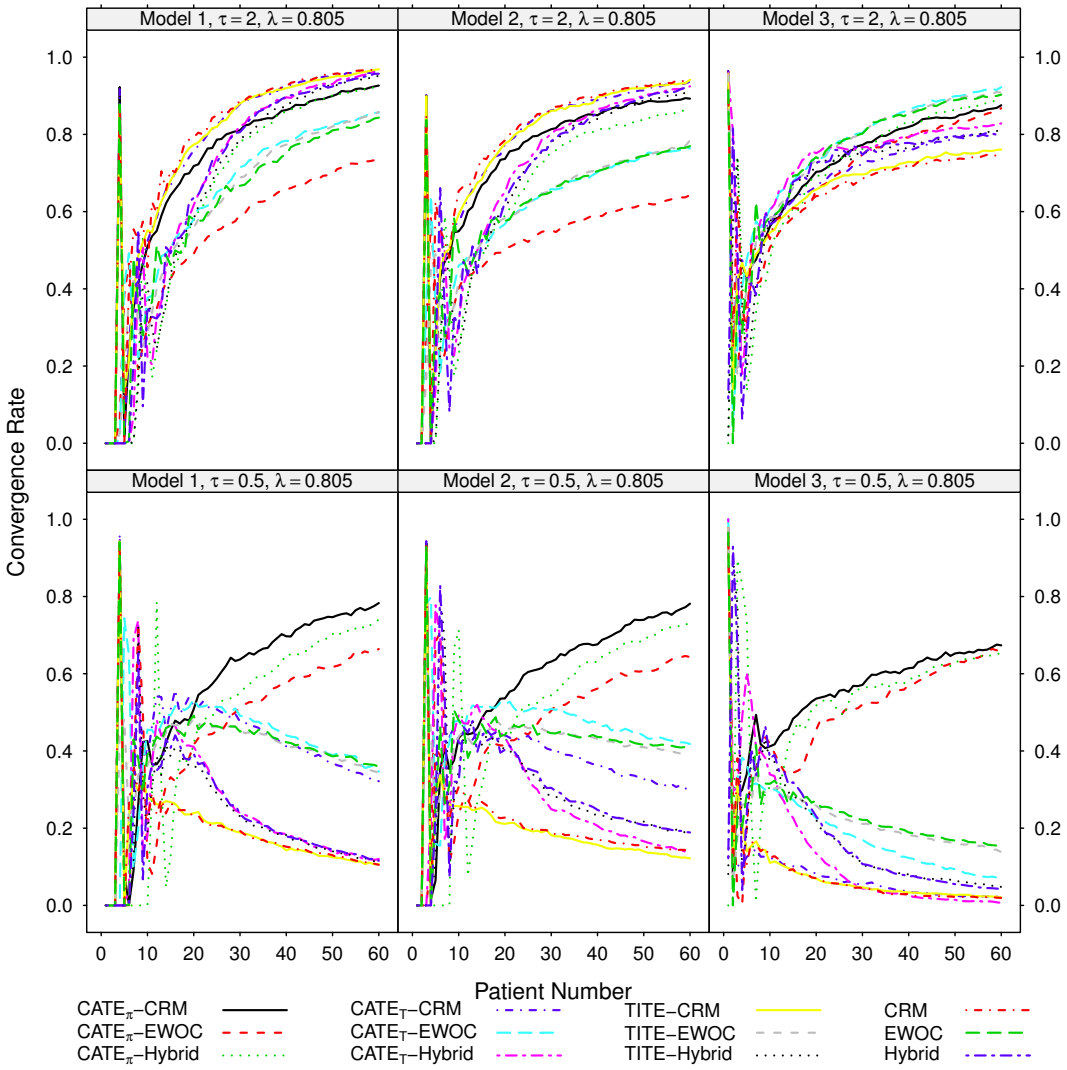


Figure 5.5: Convergence Rates with Different Methods When $\lambda = 0.805$

5.2.2 When the Planned Follow-up Time is Short ($\tau = 0.5$)

For Model 1, when the planned follow-up time is short ($\tau = 0.5$), the CATE model using MTD_π maintains relatively high proportion of recommendation at the correct dose level, while the TITE and conventional approach overestimate the MTD and recommend the overdose level (dose level 6) most of the time. The CATE model using MTD_T also recommends dose level 6 more frequently than the other doses, but to a lesser degree than TITE or conventional approach. As shown in Figure 3.2a, the probability of DLT still increases at time 0.5, and the planned follow-up time of

0.5 is too short to observe most of the DLT for dose level 5 or 6 (only 43% of all DLT expected on MTD level 5). The EWOC design has lower percentages of correct recommendation than CRM or hybrid for CATE model using MTD_π, but has lower percentage of overdose, as it recommends a lower dose level ($MTD-1$) more often than CRM or hybrid. Although TITE, conventional and CATE model using MTD_T overestimate the MTD, the EWOC design tends to overestimate to a less degree than the corresponding CRM or hybrid design, and recommends the correct MTD level (level 5) more than CRM or hybrid.

For Model 2, when the planned follow-up time is short ($\tau = 0.5$), the CATE model using MTD_π maintains relatively high proportion of recommendation at the correct dose level. The TITE and conventional approach overestimate the MTD and recommend the $MTD+1$ level (dose level 5) most of the time when the planned follow up time is relatively short. The CATE model using MTD_T also recommends dose level 5 more frequently than the other doses, but it recommends dose level 6 ($MTD+2$) lower than TITE or conventional approach. The probability of DLT in the true model continues to increase at time 0.5 as depicted in Figure 3.2b. The planned follow-up time of 0.5 is too short and only 43% of all DLT is expected on MTD level 4 by time 0.5. The EWOC design has lower percentages of correct recommendation than CRM or hybrid for CATE model using MTD_π, but has lower percentage of overdose, as it recommends a lower dose level ($MTD-1$) more often than CRM or hybrid. Although TITE, conventional and CATE model using MTD_T overestimate the MTD, the EWOC design has a higher percentages of recommendation at the correct dose level (level 4) than the corresponding CRM or hybrid design, and has a lower percentage of overdose recommendation.

For Model 3, when the planned follow-up time is short ($\tau = 0.5$), the CATE model using MTD_π has relatively high proportion of recommendation at the correct dose level. The CATE model using MTD_T, TITE and conventional approach all overestimate the MTD, and recommend the $MTD+1$ level (dose level 3) most of the time. However, the CATE model using MTD_T is less likely to recommend overdose levels above $MTD+1$ (dose level 4, 5 and 6) compared with TITE or conventional approach. As depicted in

Figure 3.2c, the planned follow-up time of 0.5 is too short and only 39% of all DLT is expected on MTD (level 2) by time 0.5. The EWOC design is comparable to CRM or hybrid in terms of correct recommendation for CATE model using MTD_π, but has lower percentage of overdose, as it recommends a lower dose level ($MTD-1$) more often than CRM or hybrid. Although TITE, conventional and CATE model using MTD_T overestimate the MTD, the EWOC design has a higher percentages of recommendation at the correct dose level (level 2) than the corresponding CRM or hybrid design, and has a lower percentage of overdose recommendation.

As showed in the lower panels of Figure 5.5, for all the models, the convergence rates of CRM, EWOC or hybrid for CATE model using MTD_π continue to increase as number of patients increases. All other methods did not converge. When the planned follow-up time is insufficient and too short, selecting doses under CATE model using the MTD_π definition is more appropriate and is better than the TITE or the conventional approaches.

5.2.3 Summary of Comparisons to Other Designs

Results in Model 1 and Model 2 are similar when the MTD level is medium (model 2) or high (model 1). When the planned follow-up time is sufficient ($\tau = 2$, 86% to 89% of all DLT expected on MTD level by time 2), the proportion of recommendation at the correct dose level for the CATE model using MTD_T (for all designs) is similar to that using the TITE approach and the conventional approach, and the percentage is lower for CATE model using MTD_π compared with other approaches. The EWOC design has lower percentages of correct recommendation than CRM or hybrid for all approaches. The percentages of recommendation at the overdose level are very small ($< 1.7\%$) in all scenarios. EWOC is less likely to recommend the overdose level than CRM or hybrid, and recommends a lower dose level ($MTD - 1$) more often than CRM or hybrid. The trend for the convergence rates is also similar. The convergence rates for all designs continue to increase as number of patients increases, but most of the gain has been obtained with 30 patients. The convergence rates for the CRM or hybrid designs are higher than the corresponding EWOC designs.

Under true model 3 with low MTD (MTD level = 2), when the planned follow-up time is sufficient (89% of all DLT expected on MTD level by time 2), the proportion of correct recommendation for EWOC design using CATE model MTD_T is slightly higher than that using the TITE approach and the conventional approach, and the proportion is lower for the CATE model using MTD_π. For CRM or hybrid design, the performance for the CATE model using MTD_π is better than MTD_T, and the CATE model (both MTD_T and MTD_π) is better than the TITE approach or the conventional approach. EWOC design is better than (in most cases) or at least comparable to CRM or hybrid (CATE model using MTD_π) in terms of percentage of correct recommendation. EWOC has lower percentages of recommendation at the overdose level than CRM or hybrid, and recommends a lower dose level ($MTD - 1$) more than CRM or hybrid. Similar to model 1 and model 2, the convergence rates for all designs continue to increase as number of patients increases, but most of the gain has been obtained with 30 patients. The convergence rates for the EWOC designs are higher than the corresponding CRM or hybrid designs in most scenarios.

When the planned follow-up time is short ($\tau = 0.5$, only 39% to 43% of all DLT expected on MTD level by time 0.5), CATE model using MTD_π is superior to other approaches for all true models considered. The CATE model using MTD_π maintains relatively high proportion of recommendation at the correct dose level. The CATE model using MTD_T, TITE and conventional approach all overestimate the MTD, and recommend the $MTD + 1$ level most of the time. However, the CATE model using MTD_T is less likely to recommend overdose levels above $MTD + 1$ compared with TITE or conventional approach. The EWOC design has lower percentages of correct recommendation than CRM or hybrid for CATE model using MTD_π under true model 1 or true model 2; while under true model 3, the EWOC design is comparable to CRM or hybrid. The EWOC design has a lower percentage of overdose recommendation than CRM or hybrid in all scenarios examined. The convergence rates for CATE model using MTD_π continue to increase as number of patients increases. All other methods do not converge.

5.3 Performance of the CATE Design by Comparing to Other Existing Designs with High Baseline Hazard Using Uniform Priors

The performance of the CATE design is evaluated using the non-informative prior on the baseline hazard rate λ in this section. Results presented here used uniform priors for α, β, γ and λ. The performance of the CATE design is compared to the TITE approach and the conventional approach under three true dose-response models with high baseline hazard (hazard rate $\lambda = 0.805$). Simulations were run for CRM, EWOC and Hybrid designs varying the planned follow-up time (follow-up time window $\tau = 2$ vs. $\tau = 0.5$). CRM2 in Chu et al. (2009) are presented as the CRM model in the results except for CATE design using MTD_T under true model 3 when $\tau = 2$, where CRM3 results are presented as CRM because of the computation problem with CRM2.

The planned follow up time of 2 is sufficient for the true dose-response models and the baseline hazard under consideration, where 86%, 89%, and 89% of all DLT on the respective MTD level would have been expected by time 2 under Model 1, Model 2, and Model 3.

The planned follow up time of 0.5 is insufficient for the true dose-response models and the baseline hazard under consideration, where 43%, 43%, and 39% of all DLT on the respective MTD level would have been expected by time 0.5 under Model 1, Model 2, and Model 3. Under these scenarios, the true MTD should be defined using the MTD_π definition.

5.3.1 Model 1

As shown in Table 5.7, when the planned follow-up time is sufficient ($\tau = 2$), the proportion of recommendation at the correct dose level for the CATE design using MTD_T for both EWOC and hybrid designs is similar to that using the TITE approach and the conventional approach; the proportion is about 12 percentage points lower for CRM design compared with TITE or conventional approach. The proportion of correct recommendation for the CATE design using MTD_π is lower than that using MTD_T for all designs. This might be explained by the fact that the probability of DLT under true

model 1 at dose level 5 is closer to the target toxicity level θ at time 2 (corresponding to the MTD_T definition) than at time 5 (corresponding to the MTD_π definition since by time 5 the probability of DLT has reached a plateau) as depicted in Figure 3.2 (a). The EWOC design has lower percentages of correct recommendation than CRM or Hybrid for most scenarios. The hybrid design performs better than CRM for CATE design using either MTD_T or MTD_π. The percentages of recommendation at the overdose level (level 6) are very small ($< 1.1\%$) in all scenarios.

When the planned follow-up time is short ($\tau = 0.5$), the CATE design using MTD_π maintains relatively high proportion of recommendation at the correct dose level for hybrid and CRM; EWOC recommends the $MTD - 1$ level most frequently. The TITE and conventional approach overestimate the MTD and recommend the $MTD + 1$ level (dose level 6) most of the time when the planned follow-up time is short. The CATE design using MTD_T for EWOC and hybrid also recommends dose level 6 more frequently than the other doses; however for CRM it has high proportion (83%) of correct recommendation. The probability of DLT in the true model continues to increase at time 0.5 as depicted in Figure 3.2 (a). The planned follow-up time of 0.5 is too short and only 43% of all DLT is expected on MTD level 5 by time 0.5. The hybrid design has higher percentages of correct recommendation than CRM or EWOC for CATE design using MTD_π. Although TITE, conventional and CATE design using MTD_T overestimate the MTD, the EWOC design tends to overestimate to a less degree than the corresponding CRM or hybrid design, and recommends the correct MTD level (level 5) more than CRM or hybrid, except in the case for CRM under CATE design using MTD_T. When the planned follow-up time is insufficient and too short, selecting doses under CATE design using the MTD_π definition is more appropriate than the MTD_T definition, and CATE design using MTD_π is better than the TITE or conventional approach.

Table 5.7: Comparison Percentage of Recommended MTD by Dose Level Under True Model 1, Exponential Four-Parameter Model with Non-Informative Prior on λ and $\lambda = 0.805$

Uniform Prior $\tau = 2$							
Dose levels	1	2	3	4	5(MTD)	6	Duration
CATE_π							
CRM	0.00	0.00	0.75	30.80	68.35	0.10	60.6
EWOC	0.00	0.30	3.25	52.00	44.40	0.05	60.7
Hybrid	0.00	0.00	0.35	21.95	77.30	0.40	60.7
CATE_T							
CRM	0.00	0.00	0.00	15.15	84.70	0.15	60.9
EWOC	0.00	0.00	0.05	12.95	86.85	0.15	60.7
Hybrid	0.00	0.00	0.00	3.10	96.10	0.80	61.0
TITE							
CRM	0.00	0.00	0.00	2.95	96.85	0.20	60.9
EWOC	0.00	0.00	0.00	13.90	85.85	0.25	60.8
Hybrid	0.00	0.00	0.00	3.80	95.15	1.05	60.8
Conventional							
CRM	0.00	0.00	0.00	2.65	96.85	0.50	111.1
EWOC	0.00	0.00	0.05	15.50	84.35	0.10	114.6
Hybrid	0.00	0.00	0.00	3.55	95.70	0.75	113.3
Uniform Prior $\tau = 0.5$							
Dose levels	1	2	3	4	5(MTD)	6	Duration
CATE_π							
CRM	0.00	0.05	2.45	37.65	56.90	2.95	59.6
EWOC	0.00	0.65	6.30	52.25	39.50	1.30	59.4
Hybrid	0.00	0.05	0.50	20.65	67.30	11.50	59.5
CATE_T							
CRM	0.00	0.00	0.00	1.30	82.70	16.00	59.9
EWOC	0.00	0.00	0.00	0.00	31.05	68.95	59.7
Hybrid	0.00	0.00	0.00	0.00	10.15	89.85	59.5
TITE							
CRM	0.00	0.00	0.00	0.00	10.70	89.30	59.3
EWOC	0.00	0.00	0.00	0.05	34.45	65.50	59.7
Hybrid	0.00	0.00	0.00	0.05	11.50	88.45	59.3
Conventional							
CRM	0.00	0.00	0.00	0.00	10.50	89.50	64.3
EWOC	0.00	0.00	0.00	0.10	35.90	64.00	64.8
Hybrid	0.00	0.00	0.00	0.00	11.55	88.45	64.5

5.3.2 Model 2

As shown in Table 5.8, when the planned follow-up time τ equals to 2, the proportion of recommendation at the correct dose level for the CATE design using MTD_T for both EWOC and hybrid designs is similar to that using the TITE approach and the conventional approach; the proportion is about 11 percentage points lower for CRM design compared with TITE or conventional approach. The proportion of correct recommendation for the CATE design using MTD_π is lower than that using MTD_T for all designs. As depicted in Figure 3.2 (b), the probability of DLT under true model 2 at dose level 4 is closer to the target toxicity level θ for the MTD_T definition (at time 2) than for the MTD_π definition. The EWOC design has lower percentages of correct recommendation than CRM or Hybrid for all approaches (CATE, TITE or conventional approach). The hybrid design performs better than CRM for CATE design using either MTD_T or MTD_π. The percentages of recommendation at the overdose level (level 5 or 6) are very small ($< 1.7\%$) in all scenarios.

When the planned follow-up time is short ($\tau = 0.5$), the CATE design using MTD_π maintains relatively high proportion of recommendation at the correct dose level for hybrid and CRM; EWOC recommend the $MTD - 1$ level most frequently. The TITE and conventional approach overestimate the MTD and recommend the overdose level (dose level 5) most of the time. The CATE design using MTD_T for EWOC and hybrid also recommends dose level 5 more frequently than the other doses, however for CRM it has relatively high proportion (67%) of correct recommendation. As shown in Figure 3.2 (b), the probability of DLT still increases at time 0.5, and the planned follow up time of 0.5 is too short to observe most of the DLT for dose level 4 (only 43% of all DLT expected on MTD level 4). The hybrid design has higher percentages of correct recommendation than CRM or EWOC for CATE design using MTD_π. Although TITE, conventional and CATE design using MTD_T overestimate the MTD, the EWOC design tends to overestimate to a less degree than the corresponding CRM or hybrid design, and recommends the correct MTD level (level 4) more than CRM or hybrid, except in the case for CRM under CATE design using MTD_T. When the planned follow-up

time is insufficient and too short, selecting doses under CATE design using the MTD_π definition is more appropriate than the MTD_T definition, and CATE design using MTD_π is better than the TITE or conventional approach.

Table 5.8: Comparison Percentage of Recommended MTD by Dose Level Under True Model 2, Exponential Four-Parameter Model with Non-Informative Prior on λ and $\lambda = 0.805$

Uniform prior $\tau = 2$							
Dose levels	1	2	3	4(MTD)	5	6	Duration
CATE$_\pi$							
CRM	0.00	1.25	32.05	66.50	0.20	0.00	60.9
EWOC	0.25	4.95	57.50	37.25	0.05	0.00	60.8
Hybrid	0.00	0.50	27.20	71.60	0.70	0.00	60.7
CATE$_T$							
CRM	0.00	0.00	16.80	82.85	0.35	0.00	60.5
EWOC	0.00	0.00	21.95	78.00	0.05	0.00	61.0
Hybrid	0.00	0.00	6.25	92.60	1.15	0.00	60.4
TITE							
CRM	0.00	0.00	4.75	94.15	1.10	0.00	60.6
EWOC	0.00	0.10	21.35	78.50	0.05	0.00	60.5
Hybrid	0.00	0.00	7.20	91.45	1.35	0.00	60.6
Conventional							
CRM	0.00	0.00	4.90	94.15	0.95	0.00	110.1
EWOC	0.00	0.00	22.85	76.95	0.20	0.00	113.6
Hybrid	0.00	0.00	6.50	91.85	1.65	0.00	112.7
Uniform prior $\tau = 0.5$							
Dose levels	1	2	3	4(MTD)	5	6	Duration
CATE$_\pi$							
CRM	0.00	2.10	29.90	64.10	3.90	0.00	59.7
EWOC	0.75	8.15	46.55	43.20	1.35	0.00	59.4
Hybrid	0.00	0.60	15.80	71.75	11.70	0.15	59.4
CATE$_T$							
CRM	0.00	0.00	0.35	66.85	32.50	0.30	59.7
EWOC	0.00	0.00	0.25	35.10	60.30	4.35	59.7
Hybrid	0.00	0.00	0.00	13.45	70.65	15.90	59.5
TITE							
CRM	0.00	0.00	0.00	12.15	65.05	22.80	59.5
EWOC	0.00	0.00	0.55	39.15	54.65	5.65	59.5
Hybrid	0.00	0.00	0.05	18.70	64.05	17.20	59.5
Conventional							
CRM	0.00	0.00	0.00	14.25	62.75	23.00	64.2
EWOC	0.00	0.00	0.65	40.60	53.10	5.65	64.7
Hybrid	0.00	0.00	0.00	18.90	62.80	18.30	64.5

5.3.3 Model 3

As shown in Table 5.9, when τ equals to 2, the proportion of recommendation at the correct dose level for the CATE design using MTD_T for EWOC design is a little better than that using the TITE approach and the conventional approach, and the proportion for the CATE design using MTD_π is lower than the TITE approach and the conventional approach. For CRM and hybrid design, the performance for the CATE design (both MTD_T and MTD_π) is better than the TITE approach or the conventional approach. The proportion of correct recommendation for the CATE design using MTD_π is higher than that using MTD_T for CRM and hybrid design, and lower for EWOC design. As shown in Figure 3.2 (c), the probability of DLT under true model 3 at dose level 2 is closer to the target toxicity level θ for the MTD_π definition than for the MTD_T definition (at time 2). The percentage of correct recommendation is higher for EWOC design than CRM or hybrid design for all approaches except for CATE model using MTD_π, where CRM or hybrid design performs better than EWOC. The hybrid design performs better than CRM for all approaches. EWOC has lower percentages of recommendation at the overdose level than CRM or hybrid.

When the planned follow-up time is short ($\tau = 0.5$), the CATE design using MTD_π has relatively high proportion of recommendation at the correct dose level. The CATE design using MTD_T, TITE and conventional approach all overestimate the MTD, and recommend the $MTD + 1$ level (dose level 3) most of the time. However, the CATE model using MTD_T is less likely to recommend overdose levels above $MTD + 1$ (dose level 4, 5 and 6) compared with TITE or conventional approach. As depicted in Figure 3.2 (c), the planned follow-up time of 0.5 is too short and only 39% of all DLT is expected on MTD (level 2) by time 0.5. The EWOC design is comparable to CRM or hybrid in terms of correct recommendation for CATE design using MTD_π, but has lower percentage of overdose, as it recommends a lower dose level ($MTD - 1$) more often than CRM or hybrid. Although TITE, conventional and CATE design using MTD_T overestimate the MTD, the EWOC design has a higher percentages of recommendation at the correct dose level (level 2) than the corresponding CRM or

hybrid design, and has a lower percentage of overdose recommendation. When the planned follow-up time is insufficient and too short, selecting doses under CATE design using the MTD_π definition is more appropriate than the MTD_T definition, and CATE design using MTD_π is better than the TITE or conventional approach.

Table 5.9: Comparison Percentage of Recommended MTD by Dose Level Under True Model 3, Exponential Four-Parameter Model with Non-Informative Prior on λ and $\lambda = 0.805$

Uniform prior $\tau = 2$							
Dose levels	1	2(MTD)	3	4	5	6	Duration
$CATE_\pi$							
CRM	6.85	87.35	5.80	0.00	0.00	0.00	60.9
EWOC	28.10	70.60	1.30	0.00	0.00	0.00	60.7
Hybrid	6.35	88.10	5.55	0.00	0.00	0.00	60.7
$CATE_T$							
CRM*	0.00	81.25	18.75	0.00	0.00	0.00	60.6
EWOC	2.60	92.30	5.10	0.00	0.00	0.00	60.8
Hybrid	0.10	82.60	17.25	0.05	0.00	0.00	60.9
TITE							
CRM	0.25	76.10	23.65	0.00	0.00	0.00	60.5
EWOC	3.75	90.80	5.45	0.00	0.00	0.00	60.5
Hybrid	0.75	80.90	18.35	0.00	0.00	0.00	60.6
Conventional							
CRM	0.95	74.45	24.55	0.05	0.00	0.00	109.3
EWOC	3.50	90.30	5.75	0.00	0.00	0.00	113.7
Hybrid	0.70	80.20	19.10	0.00	0.00	0.00	112.5
Uniform prior $\tau = 0.5$							
Dose levels	1	2(MTD)	3	4	5	6	Duration
$CATE_\pi$							
CRM	5.00	68.20	26.45	0.35	0.00	0.00	59.6
EWOC	28.30	65.15	6.50	0.05	0.00	0.00	59.6
Hybrid	3.80	63.55	31.95	0.65	0.05	0.00	59.4
$CATE_T$							
CRM	0.00	4.65	88.00	7.35	0.00	0.00	59.5
EWOC	0.00	8.25	82.25	9.50	0.00	0.00	59.6
Hybrid	0.00	0.95	70.00	28.60	0.40	0.05	59.4
TITE							
CRM	0.00	2.05	56.75	39.30	1.90	0.00	59.3
EWOC	0.00	13.80	71.75	14.40	0.05	0.00	59.3
Hybrid	0.00	4.80	68.20	26.35	0.65	0.00	59.5
Conventional							
CRM	0.00	1.95	57.80	38.55	1.70	0.00	64.3
EWOC	0.00	15.20	70.55	14.10	0.10	0.05	64.5
Hybrid	0.00	4.25	69.25	25.80	0.70	0.00	64.5

*: CRM3 results are presented.

5.3.4 Summary

When 86% to 89% of all DLT expected on MTD level by the planned follow-up time, the follow-up time ($\tau = 2$) is sufficient relative to the risk of time to toxicity. When the true MTD is above the mid-level of the dose range considered (Model 1 or Model 2), the performance for CATE design using MTD_T for both EWOC and hybrid designs is similar to that using the TITE approach or conventional approach, and is better than CATE design using MTD_π. The performance for CATE design using MTD_T is lower for CRM design compared with TITE or conventional approach, and is better than CATE design using MTD_π. The hybrid design recommends the target dose on MTD more often than CRM or EWOC for CATE design using either MTD_T or MTD_π. The choice of prior on the baseline hazard rate has little effect on the performance of the CATE design using MTD_T for EWOC or hybrid; the performance is higher for CATE design using MTD_π for all designs or MTD_T for CRM when using informative prior than the non-informative prior. When the true MTD is below the mid-level of the dose range considered (model 3), the performance for EWOC design using CATE design MTD_T is a little better than that using the TITE approach and the conventional approach, and is better than the CATE model using MTD_π. For CRM or hybrid design, the performance for the CATE design using MTD_π is better than MTD_T, and both are better than the TITE approach or the conventional approach. The percentage of correct recommendation is higher for EWOC design than CRM or hybrid design for all approaches except for CATE model using MTD_π, where CRM or hybrid design performs better than EWOC. The hybrid design performs better than CRM for all approaches. EWOC has lower percentages of recommendation at the overdose level than CRM or hybrid. The choice of prior on the baseline hazard rate has little effect on the performance of the CATE design using MTD_T or MTD_π, except in the case of EWOC using MTD_π when the performance is higher using informative prior than the non-informative prior.

When the planned follow-up time is short (39% to 43% of all DLT expected on MTD level by the planned follow-up time), CATE design using MTD_π is superior to TITE or

conventional approaches for all true models considered. TITE or conventional approach overestimates MTD and recommend the $MTD + 1$ level most frequently. When the true MTD is above the mid-level of the dose range considered (Model 1 or Model 2), the CATE design using MTD_π maintains relatively high proportion of recommendation at the correct dose level for hybrid and CRM; EWOC recommend the $MTD - 1$ level most frequently. The CATE design using MTD_T for EWOC and hybrid also recommends $MTD + 1$ level most frequently; however for CRM it has high proportion of correct recommendation. When the true MTD is below the mid-level (Model 3), the CATE design using MTD_π maintains relatively high proportion of recommendation at the correct dose level for all designs. CATE design using MTD_T overestimates MTD and recommends $MTD + 1$ level most frequently. For CATE design using MTD_π, hybrid designs have better convergence rates and recommend the target dose on MTD more often than CRM or EWOC when the true MTD is above the mid-level of the dose range considered; when the true MTD is below the mid-level, EWOC designs are comparable to CRM or hybrid. EWOC controls the overdose proportion better than CRM or hybrid. The choice of prior on the baseline hazard rate has little effect on the performance of the CATE design using MTD_π when the true MTD is below the mid-level. The performance of the CATE design using MTD_π is higher using informative prior than the non-informative prior when the true MTD is above the mid-level of the dose range considered.

5.4 Performance of the CATE Design by Comparing to Other Existing Designs with Low Baseline Hazard Using Uniform Priors

The performance of the CATE design is evaluated using the non-informative prior on the baseline hazard rate λ in this section. Results presented here used uniform priors for α, β, γ and λ. The performance of the CATE design is compared to the TITE approach and the conventional approach under three true dose-response models with low baseline hazard (hazard rate $\lambda = 0.403$). Simulations were run for CRM, EWOC and Hybrid designs when the planned follow-up time τ equals to 2. CRM2 in Chu et al. (2009) are presented as the CRM model in the results.

This represents the case where the planned follow-up time is moderate relative to the risk of time to toxicity, where 67%, 67%, and 63% of all DLT on the respective MTD level would have been expected by time 2 under Model 1, Model 2, and Model 3, respectively.

5.4.1 Model 1

As shown in Table 5.10, the proportion of recommendation at the correct dose level for the CATE design using MTD_T for both EWOC and hybrid designs is a little better than that using the TITE approach and the conventional approach; the proportion is about 10 percentage points higher for CRM design compared with TITE or conventional approach. The proportion of correct recommendation for the CATE design using MTD_π is lower than that using MTD_T for all designs. Figure 3.3 (a) shows that the probability of DLT under true model 1 at dose level 5 is below the target toxicity level θ at time 2 (correspond to MTD_T) and slightly above the target toxicity level θ at time 5 (correspond to MTD_π). The percentage of correct recommendation for EWOC is lower than CRM or hybrid for CATE design using MTD_π, while the percentage is higher than CRM or hybrid for all other approaches. The hybrid design performs better than CRM for CATE design using MTD_π, and performs slightly lower than CRM for CATE design using MTD_T. EWOC has lower percentages of recommendation at the overdose level than CRM or hybrid in most scenarios. The total trial duration using the CATE design (both MTD_π and MTD_T) is much shorter compared with the conventional approach, and is similar to the TITE approach.

Table 5.10: Comparison Percentage of Recommended MTD by Dose Level Under True Model 1, Exponential Four-Parameter Model with Non-Informative Prior on λ and $\lambda = 0.403$

Dose levels	1	2	3	4	5(MTD)	6	Duration
Uniform prior $\tau = 2$							
CATE_π							
CRM	0.00	0.00	2.05	40.45	56.70	0.80	61.0
EWOC	0.00	0.40	6.20	54.10	38.90	0.40	60.9
Hybrid	0.00	0.05	0.55	25.20	71.90	2.30	60.9
CATE_T							
CRM	0.00	0.00	0.00	10.90	87.40	1.70	60.7
EWOC	0.00	0.00	0.00	1.90	92.30	5.80	61.0
Hybrid	0.00	0.00	0.00	0.10	79.50	20.40	60.5
TITE							
CRM	0.00	0.00	0.00	0.15	76.75	23.10	60.7
EWOC	0.00	0.00	0.00	2.55	90.60	6.85	60.9
Hybrid	0.00	0.00	0.00	0.10	77.90	22.00	60.8
Conventional							
CRM	0.00	0.00	0.00	0.25	79.80	19.95	113.5
EWOC	0.00	0.00	0.00	3.05	91.40	5.55	116.6
Hybrid	0.00	0.00	0.00	0.40	77.90	21.70	115.2

5.4.2 Model 2

As shown in Table 5.11, the proportion of recommendation at the correct dose level for the CATE design using MTD_T for both EWOC and hybrid designs is a little better than that using the TITE approach and the conventional approach; the proportion is about 17 percentage points higher for CRM design compared with TITE or conventional approach. The proportion of correct recommendation for the CATE design using MTD_π is lower than that using MTD_T for all designs. Figure 3.3 (b) shows that the probability of DLT under true model 2 at dose level 4 is below the target toxicity level θ at time 2 (correspond to MTD_T) and slightly above the target toxicity level θ at time 5 (correspond to MTD_π). The percentage of correct recommendation for EWOC is lower than CRM or hybrid for CATE design using MTD_π, while the percentage is higher than CRM or hybrid for all other approaches. The hybrid design performs better than CRM for CATE design using MTD_π, and performs slightly lower than CRM for for CATE design using MTD_T. EWOC has lower percentages of recommendation

at the overdose level than CRM or hybrid in most scenarios. The total trial duration using the CATE design (both MTD_π and MTD_T) is much shorter compared with the conventional approach, and is similar to the TITE approach.

Table 5.11: Comparison Percentage of Recommended MTD by Dose Level Under True Model 2, Exponential Four-Parameter Model with Non-Informative Prior on λ and $\lambda = 0.403$

Dose levels	1	2	3	4(MTD)	5	6	Duration
Uniform prior $\tau = 2$							
CATE_π							
CRM	0.00	2.60	36.00	60.35	1.05	0.00	60.8
EWOC	1.15	8.35	48.80	41.40	0.30	0.00	60.8
Hybrid	0.00	0.35	25.30	72.00	2.35	0.00	61.1
CATE_T							
CRM	0.00	0.00	8.00	88.75	3.25	0.00	60.9
EWOC	0.00	0.00	4.55	90.05	5.40	0.00	61.0
Hybrid	0.00	0.00	0.35	79.50	20.10	0.05	60.5
TITE							
CRM	0.00	0.00	0.60	71.85	27.55	0.00	60.8
EWOC	0.00	0.00	5.70	86.90	7.35	0.05	60.9
Hybrid	0.00	0.00	1.15	75.65	23.15	0.05	60.9
Conventional							
CRM	0.00	0.00	0.80	71.75	27.35	0.10	112.0
EWOC	0.00	0.00	6.30	87.30	6.35	0.05	115.8
Hybrid	0.00	0.00	0.65	76.70	22.65	0.00	114.7

5.4.3 Model 3

As shown in Table 5.12, the CATE design using MTD_π has relatively high proportion of recommendation at the correct dose level for both CRM and hybrid designs and is superior to other approaches (CATE design using MTD_T, TITE or conventional approach); for EWOC it is a little better than other approaches. For EWOC design, the CATE design using MTD_T recommends the correct dose level (dose level 2) most of the time; the percentage is slightly low than that using MTD_π, but a little higher than TITE or conventional approach. For CRM and hybrid design, TITE and conventional approach overestimate the MTD and recommends the $MTD+1$ level most of the time. The CATE model using MTD_T for CRM design recommends the correct MTD dose

level more frequently than the other doses and is better than TITE or conventional approach. The CATE model using MTD_T for hybrid design also overestimate the MTD and recommends the $MTD+1$ dose level more frequently than the other doses. As shown in Figure 3.3 (c), the probability of DLT under true model 3 at dose level 2 is closer to the target toxicity level θ for the MTD_π definition (at time 5) than for the MTD_T definition (at time 2). The percentage of correct recommendation is higher for EWOC design than CRM or hybrid design for all approaches except for CATE design using MTD_π, where CRM or hybrid design performs better than EWOC. EWOC has lower percentages of recommendation at the overdose level than CRM or hybrid for all approaches. The total trial duration using the CATE design (both MTD_π and MTD_T) is much shorter compared with the conventional approach, and is similar to the TITE approach.

Table 5.12: Comparison Percentage of Recommended MTD by Dose Level Under True Model 3, Exponential Four-Parameter Model with Non-Informative Prior on λ and $\lambda = 0.403$

Dose levels	1	2(MTD)	3	4	5	6	Duration
Uniform prior $\tau = 2$							
CATE_π							
CRM	6.65	77.70	15.65	0.00	0.00	0.00	60.7
EWOC	29.00	66.85	4.15	0.00	0.00	0.00	60.7
Hybrid	4.90	76.60	18.50	0.00	0.00	0.00	60.9
CATE_T							
CRM	0.05	57.10	42.80	0.05	0.00	0.00	60.4
EWOC	0.20	63.70	36.10	0.00	0.00	0.00	60.7
Hybrid	0.00	36.30	63.55	0.15	0.00	0.00	60.8
TITE							
CRM	0.00	27.90	71.40	0.70	0.00	0.00	60.7
EWOC	0.55	62.40	36.80	0.25	0.00	0.00	60.8
Hybrid	0.00	35.85	63.90	0.25	0.00	0.00	60.7
Conventional							
CRM	0.00	29.70	69.45	0.85	0.00	0.00	110.3
EWOC	0.65	62.05	37.25	0.05	0.00	0.00	114.9
Hybrid	0.05	37.10	62.45	0.40	0.00	0.00	113.4

5.4.4 Summary

When 63% to 67% of all DLT expected on MTD level by the planned follow-up time, the follow-up time is moderate relative to the risk of time-to-DLT. When the true MTD is above the mid-level of the dose range considered (Model 1 and Model 2), the performance for CATE design using MTD_T for CRM is better than TITE or conventional approach; for EWOC or hybrid, it is similar to TITE or conventional approach. The proportion of correct recommendation for the CATE design using MTD_π is lower than that using MTD_T for all designs. The percentage of correct recommendation for EWOC is lower than CRM or hybrid for CATE design using MTD_π, while the percentage is higher than CRM or hybrid for all other approaches. When the true MTD is below the mid-level of the dose range considered (model 3), the CATE design using MTD_π is superior to other approaches and maintains relatively high proportion of recommendation at the correct dose level. For CRM or hybrid, TITE and the conventional approaches overestimate MTD and recommend $MTD + 1$ most of the time. For EWOC, TITE and the conventional approaches converge to the correct MTD dose but the performance is lower than CATE design using MTD_π. The performance for CATE model using MTD_T is better than TITE or conventional approach for CRM and similar to TITE or conventional for EWOC or hybrid. The percentage of correct recommendation is higher for EWOC design than CRM or hybrid design for all approaches except for CATE design using MTD_π, where CRM or hybrid design performs better than EWOC. The total trial duration using the CATE design (both MTD_π and MTD_T) is much shorter compared with the conventional approach, and is similar to the TITE approach for all true models considered.

5.5 Conclusion

It has been demonstrated by simulation that the proposed CATE model based on the proportional hazards model using exponential distribution under different priors and true dose-toxicity models has generally high performance with better percentage of correct dose recommendation and low overdose proportion. The performance of the

CATE design assuming an informative gamma prior for the baseline hazard rate is very similar to that assuming the baseline hazard rate is a constant. When the true MTD is above the mid-level of the dose range considered (model 1 and model 2), the hybrid design and CRM designs have better convergence rates and recommend the target dose on MTD more often than EWOC. Designs using the MTD_T definition have better performance than the corresponding one using the MTD_π. The choice of prior has little effect on the performance of the designs. When the true MTD is below the mid-level of the dose range considered (model 3), EWOC design has better performance than other designs in most cases, but it somewhat depends on the prior distributions and the MTD definition. EWOC and the hybrid design generally provide a better safety protection in limiting higher dose for patients than the CRM designs do. Comparison between the two MTD definitions depends on the designs and the prior distributions. EWOC using the MTD_T definition generally has better performance than the corresponding one using the MTD_π, while the opposite is true for other designs.

The performance of the proposed cure model is compared with TITE approach and the conventional approach under different planned follow-up time and true dose-toxicity models using the uniform prior distributions for α, β, γ. The simulation study shows that the total trial duration is similar for the cure model (both MTD_π and MTD_T) and the TITE approach, and is much shorter compared with the conventional approach in all scenarios. The performance of the different approaches depends somewhat on the underlying dose-toxicity models, the planned follow-up time and the dose selecting designs (CRM and hybrid vs. EWOC).

When 86% to 89% of all DLT expected on MTD level by the planned follow-up time, the follow-up time is sufficient relative to the risk of time-to-DLT. The performance for cure model using MTD_T is similar to that using the TITE approach or conventional approach, and is better than cure model using MTD_π when the true MTD is above the mid-level of the dose range considered. CRM and hybrid designs have better convergence rates and recommend the target dose on MTD more often than EWOC. When the true MTD is below the mid-level of the dose range considered (model 3), the performance for EWOC design using cure model MTD_T is similar to that using

the TITE approach and the conventional approach, and is better than the cure model using MTD_π. For CRM and hybrid design, the performance for the cure model using MTD_π is better than MTD_T, and both are better than the TITE approach or the conventional approach. EWOC designs have generally better performance than CRM or hybrid, and controls the overdose proportion better than CRM or hybrid.

When 63% to 67% of all DLT expected on MTD level by the planned follow-up time, the follow-up time is moderate relative to the risk of time-to-DLT. When the true MTD is above the mid-level of the dose range considered, the performance for CATE design using MTD_T is better than (for CRM) or similar to (for EWOC or hybrid) TITE or conventional approach; and is better than cure model using MTD_π. CRM or hybrid designs have better convergence rates and recommend the target dose on MTD more often than EWOC for CATE design using MTD_π; while the opposite is true for all other approaches. When the true MTD is below the mid-level of the dose range considered (model 3), the CATE model using MTD_π is superior to other approaches and maintains relatively high proportion of recommendation at the correct dose level. For CRM or hybrid, TITE and the conventional approaches overestimate MTD and recommend $MTD+1$ most of the time. The performance for cure model using MTD_T is better than TITE or conventional approach. The percentage of correct recommendation is higher for EWOC design than CRM or hybrid design for most scenarios.

When the planned follow-up time is short (39% to 43% of all DLT expected on MTD level by the planned follow up time), CATE model using MTD_π is superior to other approaches for all true models considered. The CATE model using MTD_π maintains relatively high proportion of recommendation at the correct dose level. TITE and conventional approach overestimate the MTD, and recommend the $MTD + 1$ level most of the time. Although the CATE model using MTD_T also overestimates the MTD, it is less likely to recommend overdose levels above $MTD + 1$ compared with TITE or conventional approach. CRM and hybrid designs have better convergence rates and recommend the target dose on MTD more often than EWOC when the true MTD is above the mid-level of the dose range considered. When the true MTD is below the mid-level, EWOC designs are comparable to CRM or hybrid. EWOC controls the

overdose proportion better than CRM or hybrid for all scenarios.

The effect of prior on the baseline hazard ratio (informative gamma prior vs. non-informative uniform prior) depends on the designs, MTD definition, true models and the length of the follow-up time. When the follow-up time is sufficient relative to the risk of time to toxicity, and the true MTD is above the mid-level of the dose range considered (Model 1 or Model 2), the choice of prior on the baseline hazard rate has little effect on the performance of the CATE design using MTD_T for EWOC or hybrid; the performance is higher for CATE design using MTD_π for all designs or MTD_T for CRM when using informative prior than the non-informative prior. When the true MTD is below the mid-level of the dose range considered (model 3), the choice of prior on the baseline hazard rate has generally little effect on the performance of the CATE design using MTD_T or MTD_π, except in the case of EWOC using MTD_π when the performance is higher using informative prior than the non-informative prior. When the planned follow-up time is short, the choice of prior on the baseline hazard rate has little effect on the performance of the CATE design using MTD_π when the true MTD is below the mid-level. The performance of the CATE design using MTD_π is higher using informative prior than the non-informative prior when the true MTD is above the mid-level of the dose range considered.

In conclusion, the proposed cure model has generally high percentage of correct dose recommendation and low overdose proportion. It significantly reduces the overall trial duration compared with the conventional approach. When the planned follow up time is moderate or too short relative to the risk of time-to-DLT, the cure model (especially MTD_π for the short follow up time) is superior to TITE or conventional approach, while TITE or conventional approach would overestimate MTD in certain situations.

Chapter 6

Performance of the CATE Design Based on the Proportional Hazards Model Using Weibull Distribution

The performance of the CATE Design based on the Proportional Hazards Model using Weibull distribution is presented in this chapter. This is a Weibull five-parameter models where $\alpha, \beta, \gamma, \lambda, v$ are estimated from data.

6.1 Performance When Data Were Generated Using Weibull Distribution

We first evaluate the performance of the CATE design when data were generated using Weibull distribution and estimated using the Weibull distribution.

6.1.1 Model 1

The percentages of recommendation at each dose level based on MTD_π and MTD_T for various designs under true model 1 (MTD level = 5) under the CATE design are summarized in Table 6.1. These results used the uniform prior for α, β, γ, and the informative gamma prior for λ and v.

Based on MTD_π

Table 6.1 shows that both CRM and Hybrid designs under the CATE model have high proportions ($> 91\%$) of recommendation at the correct MTD dose level (Level 5). The EWOC design has lower percentages of correct recommendation (74%) compared with other designs, and recommend a lower dose level (Level 4) more than the other designs. The hybrid design is comparable to CRM. The percentages of recommendation at the overdose level (Level 6) are very small ($\leq 0.25\%$) in all designs.

Based on MTD_T

Table 6.1 also shows that both CRM and Hybrid designs under the CATE model have high proportions ($> 92\%$) of recommendation at the correct MTD dose level (Level 5). The performance using MTD_T is slightly better than the corresponding ones using MTD_π. Similar to the cases using MTD_π, the EWOC design has lower percentages of correct recommendation compared with other designs, and recommend a lower dose level (Level 4) more often than the other designs. The hybrid design is comparable to CRM. The percentages of recommendation at the overdose level (Level 6) are very small ($\leq 0.3\%$) in all designs.

Comparison Between MTD_π and MTD_T

Under true model 1 with high MTD (MTD level = 5), the performance using MTD_T is a little better than the corresponding ones using MTD_π for all designs, although CRM and Hybrid have high proportions of recommendation at the correct MTD dose level than EWOC based on either MTD_π or MTD_T. The EWOC design has slightly lower percentages of correct recommendation compared with other designs, and recommend a lower dose level (Level 4) more often than the other designs. The percentages of recommendation at the overdose level (level 6) are very small in all designs.

Table 6.1: Percentage of Recommended MTD by Dose Level Under True Model 1, Weibull Five-Parameter Model

Dose levels	1	2	3	4	5(MTD)	6
$CATE_\pi$						
CRM	0.00	0.00	0.00	6.85	93.00	0.15
EWOC	0.00	0.00	0.05	25.85	74.10	0.00
Hybrid	0.00	0.00	0.00	8.45	91.30	0.25
$CATE_T$						
CRM	0.00	0.00	0.00	6.15	93.85	0.00
EWOC	0.00	0.00	0.00	22.45	77.50	0.05
Hybrid	0.00	0.00	0.00	6.80	92.90	0.30

6.1.2 Model 2

The percentages of recommendation at each dose level based on MTD_π and MTD_T for various designs under true model 2 (MTD level = 4) under the CATE design are summarized in Table 6.2. These results used the uniform prior for α, β, γ, and the informative gamma prior for λ and v.

Based on MTD_π

Table 6.2 shows that both CRM and Hybrid designs under the CATE model have high proportions ($> 85\%$) of recommendation at the correct MTD dose level (Level 5). The EWOC design has lower percentages of correct recommendation (62.7%) compared with other designs, and recommend a lower dose level (Level 3) more than the other designs. The hybrid design is comparable to CRM. The percentages of recommendation at the overdose level (Level 5 and 6) are very small ($\leq 0.15\%$) in all designs.

Based on MTD_T

Table 6.2 also shows that both CRM and Hybrid designs under the CATE model have high proportions ($\geq 88\%$) of recommendation at the correct MTD dose level (Level 4). The performance using MTD_T is slightly better than the corresponding ones using MTD_π. Similar to the cases using MTD_π, the EWOC design has lower percentages of correct recommendation compared with other designs, and recommend a lower dose level (Level 4) more often than the other designs. The hybrid design is comparable to CRM2. The percentages of recommendation at the overdose level (Level 5 or 6) are very small ($\leq 0.5\%$) in all designs.

Comparison Between MTD_π and MTD_T

Under true model 2 with moderate MTD (MTD level = 4), the performance using MTD_T is a little better than the corresponding ones using MTD_π for all designs, although CRM and Hybrid have high proportions of recommendation at the correct MTD dose level than EWOC based on either MTD_π or MTD_T. The EWOC design

has slightly lower percentages of correct recommendation compared with other designs, and recommend a lower dose level (Level 3) more often than the other designs. The percentages of recommendation at the overdose level (Level 5 or 6) are very small in all designs.

Table 6.2: Percentage of Recommended MTD by Dose Level Under True Model 2, Weibull Five-Parameter Model

Dose levels	1	2	3	4(MTD)	5	6
CATE_π						
CRM	0.00	0.00	11.70	88.30	0.00	0.00
EWOC	0.00	0.15	37.10	62.70	0.05	0.00
Hybrid	0.00	0.00	14.05	85.80	0.15	0.00
CATE_T						
CRM	0.00	0.00	10.10	89.70	0.20	0.00
EWOC	0.00	0.15	32.90	66.95	0.00	0.00
Hybrid	0.00	0.00	11.50	88.00	0.50	0.00

6.1.3 Model 3

The percentages of recommendation at each dose level based on MTD_π and MTD_T for various designs under true model 3 (MTD level = 2) under the CATE design are summarized in Table 6.3. These results used the uniform prior for α, β, γ, and the informative gamma prior for λ and v.

Based on MTD_π

Table 6.3 shows that all designs under the CATE model have high proportions ($> 88\%$) of recommendation at the correct MTD dose level (Level 2). The EWOC design has comparable percentages of correct recommendation compared with other designs, and recommend a lower dose level (Level 1) more than the other designs. The hybrid design is comparable to CRM. The percentages of recommendation at the overdose level are about 7 to 9% for CRM and Hybrid and is less than 2% for EWOC.

Based on MTD_T

Table 6.3 also shows that all designs under the CATE model have high proportions ($> 90\%$) of recommendation at the correct MTD dose level (Level 2). The performance using MTD_T is slightly better than the corresponding ones using MTD_π. Similar to the cases using MTD_π, the percentages of recommendation at the overdose level are about 7 to 9% for CRM and Hybrid and is less than 3% for EWOC.

Comparison Between MTD_π and MTD_T

Under true model 3 with low MTD (MTD level = 2), the performance using MTD_T is a little better than the corresponding ones using MTD_π for all designs. All designs have high proportions of recommendation at the correct MTD dose level based on either MTD_π or MTD_T. The percentages of recommendation at the overdose level are smaller in EWOC than either CRM or hybrid.

Table 6.3: Percentage of Recommended MTD by Dose Level Under True Model 3, Weibull Five-Parameter Model

Dose levels	1	2(MTD)	3	4	5	6
$CATE_\pi$						
CRM	1.35	89.60	9.05	0.00	0.00	0.00
EWOC	10.45	88.10	1.45	0.00	0.00	0.00
Hybrid	3.05	89.45	7.50	0.00	0.00	0.00
$CATE_T$						
CRM*	0.20	90.90	8.90	0.00	0.00	0.00
EWOC	5.65	92.05	2.30	0.00	0.00	0.00
Hybrid	1.05	91.30	7.65	0.00	0.00	0.00

*: CRM3 results are presented.

6.2 Performance When Data Were Generated Using Weibull Distribution but Estimated Using Exponential Distribution

In this section, we evaluate the performance of the CATE design when data were generated using Weibull distribution but estimated using the exponential distribution (exponential four-parameter model).

6.2.1 Model 1

The percentages of recommendation at each dose level based on MTD_π and MTD_T for various designs under true model 1 (MTD level = 5) under the CATE design are summarized in Table 6.4. These results used the uniform prior for α, β, γ, and the informative gamma prior for λ.

Based on MTD_π

Table 6.4 shows that both CRM and Hybrid designs under the CATE model have relatively high proportions ($> 79\%$) of recommendation at the correct MTD dose level (Level 5). The EWOC design has lower percentages of correct recommendation (53%) compared with other designs, and recommend a lower dose level (Level 4) more than the other designs. The hybrid design is comparable to CRM. The percentages of recommendation at the overdose level (Level 6) are very small ($\leq 0.1\%$) in all designs. However, the percentage of recommendation at the correct MTD dose level for all designs is smaller than the corresponding one estimated using the correct model (Weibull distribution) as shown in Table 6.1.

Based on MTD_T

Table 6.4 also shows that both CRM and Hybrid designs under the CATE model have high proportions ($> 92\%$) of recommendation at the correct MTD dose level (Level 5). The performance using MTD_T is better than the corresponding ones using MTD_π. Similar to the cases using MTD_π, the EWOC design has lower percentages of correct recommendation compared with other designs, and recommend a lower dose level (Level 4) more often than the other designs. The hybrid design is comparable to CRM. The percentages of recommendation at the overdose level (Level 6) are very small ($\leq 0.2\%$) in all designs. Unlike the case using MTD_π, the percentages of recommendation at the correct MTD dose level for all designs are comparable to the corresponding one estimated using the correct model (Weibull distribution) as shown in Table 6.1.

Comparison Between MTD_π and MTD_T

Under true model 1 with high MTD (MTD level = 5), the performance using MTD_T is better than the corresponding ones using MTD_π for all designs, although CRM and Hybrid have high proportions of recommendation at the correct MTD dose level than EWOC based on either MTD_π or MTD_T. The EWOC design has lower percentages of correct recommendation compared with other designs, and recommend a lower dose level (Level 4) more often than the other designs. The percentages of recommendation at the overdose level (level 6) are very small in all designs. The percentages of recommendation at the correct MTD dose level for all designs are lower than the corresponding one estimated using the correct model (Weibull distribution) based on MTD_π; but comparable to the corresponding one estimated using the correct model based on MTD_T.

Table 6.4: Percentage of Recommended MTD by Dose Level Under True Model 1, Exponential Four-Parameter Model When Data Were Generated Using Weibull

Dose levels	1	2	3	4	5(MTD)	6
$CATE_\pi$						
CRM	0.00	0.00	0.00	20.50	79.40	0.10
EWOC	0.00	0.00	0.05	46.60	53.25	0.10
Hybrid	0.00	0.00	0.00	20.10	79.80	0.10
$CATE_T$						
CRM	0.00	0.00	0.00	7.15	92.70	0.15
EWOC	0.00	0.00	0.00	20.30	79.55	0.15
Hybrid	0.00	0.00	0.00	5.80	94.00	0.20

6.2.2 Model 2

The percentages of recommendation at each dose level based on MTD_π and MTD_T for various designs under true model 2 (MTD level = 4) under the CATE design are summarized in Table 6.5. These results used the uniform prior for α, β, γ, and the informative gamma prior for λ.

Based on MTD_π

Table 6.5 shows that both CRM and Hybrid designs under the CATE model have relatively high proportions ($> 71\%$) of recommendation at the correct MTD dose level (Level 4), although these percentages are lower than the corresponding ones estimated using the correct model (Weibull distribution) as shown in Table 6.2. The EWOC design has lower percentages of correct recommendation (44%) compared with other designs, and recommend a lower dose level (Level 3) more than the other designs. The hybrid design is comparable to CRM. The percentages of recommendation at the overdose level (Level 5 and 6) are very small ($\leq 0.25\%$) in all designs.

Based on MTD_T

Table 6.5 also shows that both CRM and Hybrid designs under the CATE model have high proportions ($> 85\%$) of recommendation at the correct MTD dose level (Level 4). The performance using MTD_T is better than the corresponding ones using MTD_π. Similar to the cases using MTD_π, the EWOC design has lower percentages of correct recommendation compared with other designs, and recommend a lower dose level (Level 4) more often than the other designs. The hybrid design is comparable to CRM. The percentages of recommendation at the overdose level (Level 5 or 6) are very small ($\leq 0.15\%$) in all designs. The percentages of recommendation at the correct MTD dose level for all designs are comparable to the corresponding one estimated using the correct model (Weibull distribution) as shown in Table 6.2.

Comparison Between MTD_π and MTD_T

Under true model 2 with moderate MTD (MTD level = 4), the performance using MTD_T is better than the corresponding ones using MTD_π for all designs. CRM and Hybrid have relatively high proportions of recommendation at the correct MTD dose level than EWOC based on either MTD_π or MTD_T. The EWOC design has lower percentages of correct recommendation compared with other designs, and recommend a lower dose level (Level 3) more often than the other designs. The percentages of

recommendation at the overdose level (Level 5 or 6) are very small in all designs. The percentages of recommendation at the correct MTD dose level for all designs are lower than the corresponding one estimated using the correct model (Weibull distribution) based on MTD_π; but comparable to the corresponding one estimated using the correct model based on MTD_T.

Table 6.5: Percentage of Recommended MTD by Dose Level Under True Model 2, Exponential Four-Parameter Model When Data Were Generated Using Weibull

Dose levels	1	2	3	4(MTD)	5	6
CATE_π						
CRM	0.00	0.00	25.90	74.05	0.05	0.00
EWOC	0.00	0.50	55.95	43.55	0.00	0.00
Hybrid	0.00	0.00	28.05	71.70	0.25	0.00
CATE_T						
CRM	0.00	0.00	11.85	88.05	0.10	0.00
EWOC	0.00	0.05	32.55	67.30	0.10	0.00
Hybrid	0.00	0.00	14.15	85.70	0.15	0.00

6.2.3 Model 3

The percentages of recommendation at each dose level based on MTD_π and MTD_T for various designs under true model 3 (MTD level = 2) under the CATE design are summarized in Table 6.6. These results used the uniform prior for α, β, γ, and the informative gamma prior for λ.

Based on MTD_π

Table 6.6 shows that that both CRM and Hybrid designs under the CATE model have high proportions ($> 91\%$) of recommendation at the correct MTD dose level (Level 2), and these percentages are comparable to the corresponding ones estimated using the correct model (Weibull distribution) as shown in Table 6.3. The EWOC design has lower percentages of correct recommendation compared with other designs, and recommend a lower dose level (Level 1) more than the other designs. The hybrid design is comparable to CRM. The percentage of recommendation at the overdose level is very

small ((0.3%) for EWOC, and low (about 2%) for CRM and Hybrid.

Based on MTD_T

Table 6.6 also shows that all designs under the CATE model have high proportions ($> 91\%$) of recommendation at the correct MTD dose level (Level 2). The performance using MTD_T is better than the corresponding ones using MTD_π. The hybrid design is comparable to CRM. The percentages of recommendation at the overdose level are about 5 to 6% for CRM and Hybrid and is less than 2% for EWOC. The percentages of recommendation at the correct MTD dose level for all designs are comparable to the corresponding one estimated using the correct model (Weibull distribution) as shown in Table 6.3.

Comparison Between MTD_π and MTD_T

Under true model 3 with low MTD (MTD level = 2), the performance using MTD_T is better than the corresponding ones using MTD_π for all designs. CRM and Hybrid have high proportions of recommendation at the correct MTD dose level and are better than EWOC based on MTD_π, but are similar to EWOC based on MTD_T. The percentages of recommendation at the overdose level are smaller in EWOC than either CRM or hybrid. The percentages of recommendation at the correct MTD dose level for all designs based on either MTD_π or MTD_T are comparable to the corresponding one estimated using the correct model (Weibull distribution), except for EWOC based on MTD_π where the percentage is lower than the corresponding one estimated using the correct model.

Table 6.6: Percentage of Recommended MTD by Dose Level Under True Model 3, Exponential Four-Parameter Model When Data Were Generated Using Weibull

Dose levels	1	2(MTD)	3	4	5	6
$CATE_\pi$						
CRM	5.65	92.25	2.10	0.00	0.00	0.00
EWOC	23.45	76.25	0.30	0.00	0.00	0.00
Hybrid	6.25	91.80	1.95	0.00	0.00	0.00
$CATE_T$						
CRM*	0.50	93.00	6.50	0.00	0.00	0.00
EWOC	7.25	91.55	1.20	0.00	0.00	0.00
Hybrid	1.20	93.55	5.25	0.00	0.00	0.00

*: CRM3 results are presented.

6.3 Performance When Data Were Generated Using Exponential Distribution but Estimated Using Weibull Distribution

In this section, we evaluate the performance of the CATE design when data were generated using exponential distribution but estimated using the Weibull distribution.

6.3.1 Model 1

The percentages of recommendation at each dose level based on MTD_π and MTD_T for various designs under true model 1 (MTD level = 5) under the CATE design are summarized in Table 6.7. These results used the uniform prior for α, β, γ, and the informative gamma prior for λ and v.

Based on MTD_π

Table 6.7 shows that both CRM and Hybrid designs under the CATE model have high proportions ($> 93\%$) of recommendation at the correct MTD dose level (Level 5). The EWOC design has lower percentages of correct recommendation (78%) compared with other designs, and recommend a lower dose level (Level 4) more than the other designs. The hybrid design is comparable to CRM. The percentages of recommendation at the overdose level (Level 6) are very small ($\leq 0.55\%$) in all designs. The percentages of recommendation at the correct MTD dose level for all designs are comparable or even a

little better than the corresponding one estimated using the correct model (exponential distribution) as shown in Table 5.1.

Based on MTD_T

Table 6.7 also shows that both CRM and Hybrid designs under the CATE model have high proportions ($> 95\%$) of recommendation at the correct MTD dose level (Level 5). The performance using MTD_T is better than the corresponding ones using MTD_π. The EWOC design has lower percentages of correct recommendation compared with other designs, and recommend a lower dose level (Level 4) more often than the other designs. The hybrid design is comparable to CRM. The percentages of recommendation at the overdose level (Level 6) are very small ($\leq 0.9\%$) in all designs. The percentages of recommendation at the correct MTD dose level for all designs are comparable to the corresponding one estimated using the correct model (exponential distribution) as shown in Table 5.1.

Comparison Between MTD_π and MTD_T

Under true model 1 with high MTD (MTD level = 5), the performance using MTD_T is better than the corresponding ones using MTD_π for all designs, although CRM and Hybrid have high proportions of recommendation at the correct MTD dose level than EWOC based on either MTD_π or MTD_T. The EWOC design has slightly lower percentages of correct recommendation compared with other designs, and recommend a lower dose level (Level 4) more often than the other designs. The percentages of recommendation at the overdose level (level 6) are very small in all designs. The percentages of recommendation at the correct MTD dose level for all designs are comparable to the corresponding one estimated using the correct model (exponential distribution) based on either MTD_π or MTD_T.

Table 6.7: Percentage of Recommended MTD by Dose Level Under True Model 1, Weibull Five-Parameter Model When Data Were Generated Using Exponential

Dose levels	1	2	3	4	5(MTD)	6
CATE_π						
CRM	0.00	0.00	0.00	5.85	94.05	0.10
EWOC	0.00	0.00	0.00	21.60	78.35	0.05
Hybrid	0.00	0.00	0.00	5.50	93.95	0.55
CATE_T						
CRM	0.00	0.00	0.00	3.65	95.80	0.55
EWOC	0.00	0.00	0.00	12.60	87.25	0.15
Hybrid	0.00	0.00	0.00	3.55	95.55	0.90

6.3.2 Model 2

The percentages of recommendation at each dose level based on MTD_π and MTD_T for various designs under true model 2 (MTD level = 4) under the CATE design are summarized in Table 6.8. These results used the uniform prior for α, β, γ, and the informative gamma prior for λ and v.

Based on MTD_π

Table 6.8 shows that both CRM and Hybrid designs under the CATE model have high proportions ($> 89\%$) of recommendation at the correct MTD dose level (Level 5). The EWOC design has lower percentages of correct recommendation (69%) compared with other designs, and recommend a lower dose level (Level 3) more than the other designs. The hybrid design is comparable to CRM. The percentages of recommendation at the overdose level (Level 5 and 6) are very small ($\leq 0.75\%$) in all designs. The percentages of recommendation at the correct MTD dose level for all designs are comparable to the corresponding one estimated using the correct model (exponential distribution) as shown in Table 5.1.

Based on MTD_T

Table 6.8 also shows that both CRM and Hybrid designs under the CATE model have high proportions ($> 92\%$) of recommendation at the correct MTD dose level (Level 4).

The performance using MTD_T is better than the corresponding ones using MTD_π. Similar to the cases using MTD_π, the EWOC design has lower percentages of correct recommendation compared with other designs, and recommend a lower dose level (Level 3) more often than the other designs. The hybrid design is comparable to CRM. The percentages of recommendation at the overdose level (Level 5 or 6) are very small ($\leq 1.25\%$) in all designs. The percentages of recommendation at the correct MTD dose level for all designs are comparable to the corresponding one estimated using the correct model (exponential distribution) as shown in Table 5.1.

Comparison Between MTD_π and MTD_T

Under true model 2 with moderate MTD (MTD level = 4), the performance using MTD_T is a little better than the corresponding ones using MTD_π for all designs, although CRM and Hybrid have high proportions of recommendation at the correct MTD dose level than EWOC based on either MTD_π or MTD_T. The EWOC design has slightly lower percentages of correct recommendation compared with other designs, and recommend a lower dose level (Level 3) more often than the other designs. The percentages of recommendation at the overdose level (Level 5 or 6) are very small in all designs. The percentages of recommendation at the correct MTD dose level for all designs are comparable to the corresponding one estimated using the correct model (exponential distribution) based on either MTD_π or MTD_T.

Table 6.8: Percentage of Recommended MTD by Dose Level Under True Model 2, Weibull Five-Parameter Model When Data Were Generated Using Exponential

Dose levels	1	2	3	4(MTD)	5	6
$CATE_\pi$						
CRM	0.00	0.00	10.35	89.35	0.30	0.00
EWOC	0.00	0.00	30.85	69.10	0.05	0.00
Hybrid	0.00	0.00	9.55	89.70	0.75	0.00
$CATE_T$						
CRM	0.00	0.00	5.10	94.05	0.85	0.00
EWOC	0.00	0.05	20.55	79.25	0.15	0.00
Hybrid	0.00	0.00	6.65	92.10	1.25	0.00

6.3.3 Model 3

The percentages of recommendation at each dose level based on MTD_π and MTD_T for various designs under true model 3 (MTD level = 2) under the CATE design are summarized in Table 6.9. These results used the uniform prior for α, β, γ, and the informative gamma prior for λ and v.

Based on MTD_π

Table 6.9 shows that all designs under the CATE model have high proportions ($> 86\%$) of recommendation at the correct MTD dose level (Level 2). The hybrid design is comparable to CRM. The percentages of recommendation at the overdose level are about 9 to 11% for CRM and Hybrid and is less than 2% for EWOC. The percentages of recommendation at the correct MTD dose level for all designs are comparable to the corresponding one estimated using the correct model (exponential distribution) as shown in Table 5.1.

Based on MTD_T

Table 6.9 also shows that EWOC under the CATE model has higher proportions (91.5%) of recommendation at the correct MTD dose level (Level 2) than CRM or Hybrid. The hybrid design is comparable to CRM. The performance using MTD_π is better than the corresponding ones using MTD_T for CRM and Hybrid. EWOC using MTD_T performs better than that using MTD_π. The percentages of recommendation at the overdose level are about 15 to 19% for CRM and Hybrid and is about 5% for EWOC. The performances are comparable to the corresponding ones estimated using the correct model (exponential distribution) as shown in Table 5.1.

Comparison Between MTD_π and MTD_T

Under true model 3 with low MTD (MTD level = 2), the performance using MTD_π is a little better than the corresponding ones using MTD_T for CRM and Hybrid. EWOC using MTD_T performs better than that using MTD_π. EWOC has higher proportion

of correct recommendation than CRM or Hybrid based on MTD_T; and has comparable high proportion of correct recommendation as CRM or Hybrid based on MTD_π. The percentages of recommendation at the overdose level are smaller in EWOC than either CRM or hybrid. The percentages of recommendation at the correct MTD dose level for all designs based on either MTD_π or MTD_T are comparable to the corresponding one estimated using the correct model.

Table 6.9: Percentage of Recommended MTD by Dose Level Under True Model 3, Weibull Five-Parameter Model When Data Were Generated Using Exponential

Dose levels	1	2(MTD)	3	4	5	6
CATE_π						
CRM	1.70	87.55	10.75	0.00	0.00	0.00
EWOC	11.45	86.90	1.65	0.00	0.00	0.00
Hybrid	2.35	88.40	9.25	0.00	0.00	0.00
CATE_T						
CRM*	0.20	80.50	19.30	0.00	0.00	0.00
EWOC	3.20	91.50	5.30	0.00	0.00	0.00
Hybrid	0.35	83.65	15.95	0.05	0.00	0.00

*: CRM3 results are presented.

6.4 Summary

It has been demonstrated by simulation that the proposed CATE design based on the proportional hazards model using Weibull distribution under different true dose-toxicity models has generally high performance with better percentage of correct dose recommendation and low overdose proportion. Designs using the MTD_T definition have better performance than the corresponding one using the MTD_π. When the true MTD is above the mid-level of the dose range considered (model 1 and model 2), the hybrid design and CRM designs have better convergence rates and recommend the target dose on MTD more often than EWOC. When the true MTD is below the mid-level of the dose range considered (model 3), EWOC design has comparable high performance as other designs. EWOC and the hybrid design generally provide a better safety protection in limiting higher dose for patients than the CRM designs do.

The performance of the proposed CATE model is evaluated when data were generated using Weibull distribution but estimated using the exponential distribution (exponential four-parameter model). It has been demonstrated by simulation that the proposed CATE design maintains the generally high performance with better percentage of correct dose recommendation and low overdose proportion even under the inadequate model (exponential distribution). Designs using the MTD_T definition have better performance than the corresponding one using the MTD_π. When the true MTD is above the mid-level of the dose range considered (model 1 and model 2), the hybrid design and CRM designs have better convergence rates and recommend the target dose on MTD more often than EWOC. The percentages of recommendation at the correct MTD dose level for all designs are lower than the corresponding one estimated using the correct model (Weibull distribution) based on MTD_π; but comparable to the corresponding one estimated using the correct model based on MTD_T. When the true MTD is below the mid-level of the dose range considered (model 3), the hybrid design and CRM designs have better convergence rates and recommend the target dose on MTD more often than EWOC based on MTD_π; but EWOC design has comparable high performance as other designs based on MTD_T. EWOC and the hybrid design generally provide a better safety protection in limiting higher dose for patients than the CRM designs do. The percentages of recommendation at the correct MTD dose level for all designs based on either MTD_π or MTD_T are comparable to the corresponding one estimated using the correct model, except for EWOC based on MTD_π where the percentage is lower than the corresponding one estimated using the correct model.

The performance of the proposed CATE model is also evaluated when data were generated using exponential distribution but estimated using the Weibull distribution. It has been demonstrated by simulation that the proposed CATE design maintains the generally high performance with better percentage of correct dose recommendation and low overdose proportion under this scenario. The percentages of recommendation at the correct MTD dose level for all designs based on either MTD_π or MTD_T are comparable to the corresponding one estimated using the correct model. When the true MTD is above the mid-level of the dose range considered (model 1 and model

2), the hybrid design and CRM designs have better convergence rates and recommend the target dose on MTD more often than EWOC. Designs using the MTD_T definition have better performance than the corresponding one using the MTD_π. When the true MTD is below the mid-level of the dose range considered (model 3), EWOC design has better performance than CRM or Hybrid designs based on MTD_T, and is comparable to other designs based on MTD_π. EWOC and the hybrid design generally provide a better safety protection in limiting higher dose for patients than the CRM designs do. Comparison between the two MTD definitions depends on the designs. EWOC using the MTD_T definition has better performance than the corresponding one using the MTD_π, while the opposite is true for other designs.

Chapter 7

Performance of the CATE Design Based on the Accelerated Failure Time Model Using Weibull Distribution

The performance of the CATE Design based on the Accelerated Failure Time Model using Weibull distribution is presented in this chapter. This is a Weibull five-parameter model where $\alpha, \beta, \gamma_0, \gamma_1, \sigma$ are estimated from data.

7.1 Model 1

The percentages of recommendation at each dose level based on MTD_π and MTD_T for various designs under true model 1 (MTD level = 5) under the CATE design are summarized in Table 7.1. These results used the uniform prior for α, β, γ_0, γ_1, and the inverse-gamma prior for σ.

Based on MTD_π

Table 7.1 shows that both CRM and Hybrid designs under the CATE model have high proportions ($> 91\%$) of recommendation at the correct MTD dose level (Level 5). The EWOC design has lower percentages of correct recommendation (75%) compared with other designs, and recommend a lower dose level (Level 4) more than the other designs. The hybrid design is comparable to CRM. The percentages of recommendation at the overdose level (Level 6) are very small ($\leq 0.3\%$) in all designs. These results are similar to those based on the proportional hazards model using Weibull distribution as shown in Table 6.1.

Based on MTD_T

Table 7.1 also shows that both CRM and Hybrid designs under the CATE model have high proportions ($> 92\%$) of recommendation at the correct MTD dose level (Level 5). The performance using MTD_T is a little better than the corresponding ones using MTD_π. Similar to the cases using MTD_π, the EWOC design has lower percentages of correct recommendation compared with other designs, and recommend a lower dose level (Level 4) more often than the other designs. The hybrid design is comparable to CRM. The percentages of recommendation at the overdose level (Level 6) are very small ($\leq 0.45\%$) in all designs. These results are similar to those based on the proportional hazards model using Weibull distribution as shown in Table 6.1.

Comparison Between MTD_π and MTD_T

Under true model 1 with high MTD (MTD level = 5), the performance using MTD_T is a little better than the corresponding ones using MTD_π for all designs, although CRM and Hybrid have high proportions of recommendation at the correct MTD dose level than EWOC based on either MTD_π or MTD_T. The EWOC design has slightly lower percentages of correct recommendation compared with other designs, and recommend a lower dose level (Level 4) more often than the other designs. The percentages of recommendation at the overdose level (level 6) are very small in all designs. These results are similar to those based on the proportional hazards model using Weibull distribution.

Table 7.1: Percentage of Recommended MTD by Dose Level Under True Model 1, Weibull AFT Model

Dose levels	1	2	3	4	5(MTD)	6
$CATE_\pi$						
CRM	0.00	0.00	0.00	6.50	93.50	0.00
EWOC	0.00	0.00	0.05	24.80	75.15	0.00
Hybrid	0.00	0.00	0.00	7.85	91.85	0.30
$CATE_T$						
CRM	0.00	0.00	0.00	4.55	95.40	0.05
EWOC	0.00	0.00	0.00	23.05	76.95	0.00
Hybrid	0.00	0.00	0.00	6.85	92.70	0.45

7.2 Model 2

The percentages of recommendation at each dose level based on MTD_π and MTD_T for various designs under true model 2 (MTD level = 4) under the CATE design are summarized in Table 7.2. These results used the uniform prior for α, β, γ_0, γ_1, and the inverse-gamma prior for σ.

Based on MTD_π

Table 7.2 shows that both CRM and Hybrid designs under the CATE model have high proportions ($> 83\%$) of recommendation at the correct MTD dose level (Level 5). The EWOC design has lower percentages of correct recommendation (62%) compared with other designs, and recommend a lower dose level (Level 3) more than the other designs. The hybrid design is comparable to CRM. The percentages of recommendation at the overdose level (Level 5 and 6) are very small ($\leq 0.20\%$) in all designs. These results are similar to those based on the proportional hazards model using Weibull distribution as shown in Table 6.2.

Based on MTD_T

Table 7.2 also shows that both CRM and Hybrid designs under the CATE model have high proportions ($> 87\%$) of recommendation at the correct MTD dose level (Level 4). The performance using MTD_T is a little better than the corresponding ones using MTD_π. Similar to the cases using MTD_π, the EWOC design has lower percentages of correct recommendation compared with other designs, and recommend a lower dose level (Level 4) more often than the other designs. The hybrid design is comparable to CRM. The percentages of recommendation at the overdose level (Level 5 or 6) are very small ($\leq 0.5\%$) in all designs. These results are similar to those based on the proportional hazards model using Weibull distribution as shown in Table 6.2.

Comparison Between MTD_π and MTD_T

Under true model 2 with moderate MTD (MTD level = 4), the performance using MTD_T is a little better than the corresponding ones using MTD_π for all designs, although CRM and Hybrid have high proportions of recommendation at the correct MTD dose level than EWOC based on either MTD_π or MTD_T. The EWOC design has slightly lower percentages of correct recommendation compared with other designs, and recommend a lower dose level (Level 3) more often than the other designs. The percentages of recommendation at the overdose level (Level 5 or 6) are very small in all designs. Results are similar to those based on the proportional hazards model using Weibull distribution.

Table 7.2: Percentage of Recommended MTD by Dose Level Under True Model 2, Weibull AFT Model

Dose levels	1	2	3	4(MTD)	5	6
$CATE_\pi$						
CRM	0.00	0.05	11.50	88.25	0.20	0.00
EWOC	0.00	0.20	38.15	61.60	0.05	0.00
Hybrid	0.00	0.00	16.45	83.35	0.20	0.00
$CATE_T$						
CRM	0.00	0.00	9.95	89.80	0.25	0.00
EWOC	0.00	0.05	30.60	69.30	0.05	0.00
Hybrid	0.00	0.00	12.25	87.25	0.50	0.00

7.3 Model 3

The percentages of recommendation at each dose level based on MTD_π and MTD_T for various designs under true model 3 (MTD level = 2) under the CATE design are summarized in Table 7.3. These results used the uniform prior for α, β, γ_0, γ_1, and the inverse-gamma prior for σ.

Based on MTD_π

Table 7.3 shows that all designs under the CATE model have high proportions ($> 88\%$) of recommendation at the correct MTD dose level (Level 2). The EWOC design has

comparable percentages of correct recommendation compared with other designs, and recommend a lower dose level (Level 1) more than the other designs. The hybrid design is comparable to CRM. The percentages of recommendation at the overdose level are about 7% for CRM and Hybrid and is less than 2% for EWOC. These results are similar to those based on the proportional hazards model using Weibull distribution as shown in Table 6.3.

Based on MTD_T

Table 7.3 also shows that all designs under the CATE model have high proportions ($> 90\%$) of recommendation at the correct MTD dose level (Level 2). The performance using MTD_T is a little better than the corresponding ones using MTD_π. Similar to the cases using MTD_π, the percentages of recommendation at the overdose level are about 7 to 9% for CRM and Hybrid and is less than 2% for EWOC. These results are similar to those based on the proportional hazards model using Weibull distribution as shown in Table 6.3.

Comparison Between MTD_π and MTD_T

Under true model 3 with low MTD (MTD level = 2), the performance using MTD_T is a little better than the corresponding ones using MTD_π for all designs. All designs have high proportions of recommendation at the correct MTD dose level based on either MTD_π or MTD_T. The percentages of recommendation at the overdose level are smaller in EWOC than either CRM or hybrid. Results are similar to those based on the proportional hazards model using Weibull distribution.

Table 7.3: Percentage of Recommended MTD by Dose Level Under True Model 3, Weibull AFT Model

Dose levels	1	2(MTD)	3	4	5	6
CATE_π						
CRM	1.75	90.90	7.35	0.00	0.00	0.00
EWOC	9.60	88.95	1.45	0.00	0.00	0.00
Hybrid	2.40	90.05	7.55	0.00	0.00	0.00
CATE_T						
CRM*	0.20	90.90	8.90	0.00	0.00	0.00
EWOC	7.05	91.55	1.40	0.00	0.00	0.00
Hybrid	1.60	90.60	7.80	0.00	0.00	0.00

*: CRM3 results are presented.

7.4 Summary

It has been demonstrated by simulation that the proposed CATE design based on the accelerated failure time model using Weibull distribution under different true dose-toxicity models has generally high performance with better percentage of correct dose recommendation and low overdose proportion. Designs using the MTD_T definition have better performance than the corresponding one using the MTD_π. When the true MTD is above the mid-level of the dose range considered (model 1 and model 2), the hybrid design and CRM designs have better convergence rates and recommend the target dose on MTD more often than EWOC. When the true MTD is below the mid-level of the dose range considered (model 3), EWOC design has comparable high performance as other designs. EWOC and the hybrid design generally provide a better safety protection in limiting higher dose for patients than the CRM designs do. Results are similar to those based on the proportional hazards model using Weibull distribution.

Chapter 8

Conclusions and Future Work

8.1 Conclusions

This dissertation has examined the performance of the CATE design using a Bayesian approach to find the MTD in phase I cancer clinical trials. The performance of the CATE design was evaluated using the proportional hazard model approach and the accelerated failure time model approach assuming that the DLT-free survival time for the susceptible patients follow an exponential or Weibull distribution. The robustness of the CATE design was examined using different choice of prior distribution and different dose-toxicity relationships. The performance of the CATE design was also compared to the existing methods (CRM, EWOC, Hybrid, TITE-CRM, TITE-EWOC, and TITE-hybrid).

It has been demonstrated by simulation that the proposed CATE design under different priors and true dose-toxicity models has generally high performance with high percentage of correct dose recommendation and low overdose proportion.

Under the proportional hazard model approach assuming that the DLT-free survival time for the susceptible patients follow an exponential distribution, the performance of the CATE design is very similar either assuming the baseline hazard rate is a constant or an informative gamma prior is assumed for the baseline hazard rate. When the true MTD is above the mid-level of the dose range considered (model 1 and model 2, the cases where the testing treatment is less toxic), the hybrid design and CRM designs have better convergence rates and recommend the target dose on MTD more often than EWOC. Designs using the MTD_T definition have better performance than the corresponding one using the MTD_π. The choice of prior has little effect on the performance

of the designs. When the true MTD is below the mid-level of the dose range considered (model 3, the cases where the testing treatment is more toxic), EWOC design has better performance than other designs in most cases, but it somewhat depends on the prior distributions and the MTD definition. EWOC and the hybrid design generally provide a better safety protection in limiting higher dose for patients than the CRM designs do. Comparison between the two MTD definitions depends on the designs and the prior distributions. EWOC using the MTD_T definition generally have better performance than the corresponding one using the MTD_π, while the opposite is true for other designs.

The performance of the proposed CATE design is compared with TITE approach and the conventional approach under different planned follow up time and true dose-toxicity models using the uniform prior distributions for α, β, γ. The simulation study shows that the total trial duration is similar for the CATE model (both MTD_π and MTD_T) and the TITE approach, and is much shorter compared with the conventional approach in all scenarios. The performance of the different approaches depends somewhat on the underlying dose-toxicity models, the planned follow up time and the dose selecting designs (CRM and hybrid vs. EWOC).

When 86% to 89% of all DLT expected on MTD level by the planned follow-up time, the follow-up time is sufficient relative to the risk of time to toxicity. The performance for cure model using MTD_T is similar to that using the TITE approach or conventional approach, and is better than cure model using MTD_π when the true MTD is above the mid-level of the dose range considered. CRM and hybrid designs have better convergence rates and recommend the target dose on MTD more often than EWOC. When the true MTD is below the mid-level of the dose range considered (model 3), the performance for EWOC design using cure model MTD_T is similar to that using the TITE approach and the conventional approach, and is better than the cure model using MTD_π. For CRM and hybrid design, the performance for the cure model using MTD_T is better than MTD_π, and both are better than the TITE approach or the conventional approach. EWOC design has generally better performance than CRM or hybrid, and controls the overdose proportion better than CRM or hybrid.

When 63% to 67% of all DLT expected on MTD level by the planned follow-up time, the follow-up time is moderate relative to the risk of time to toxicity. When the true MTD is above the mid-level of the dose range considered, the performance for cure model using MTD_T is better than TITE or conventional approach. For CRM and hybrid, the cure model using MTD_π has even better performance than that using MTD_T, TITE or conventional approach. CRM and hybrid designs have better convergence rates and recommend the target dose on MTD more often than EWOC in most cases. When the true MTD is below the mid-level of the dose range considered (model 3), the cure model using MTD_π is superior to other approaches and maintains high proportion of recommendation at the correct dose level. For CRM and hybrid, TITE and the conventional approaches overestimate MTD and recommend $MTD + 1$ most of the time. The performance for cure model using MTD_T is better than TITE or conventional approach. The percentage of correct recommendation is higher for EWOC design than CRM or hybrid design for all approaches.

When the planned follow-up time is short (39% to 43% of all DLT expected on MTD level by the planned follow-up time), cure model using MTD_π is superior to other approaches for all true models considered. The cure model using MTD_π maintains relatively high proportion of recommendation at the correct dose level. TITE and conventional approach overestimate the MTD, and recommend the $MTD + 1$ level most of the time. Although the cure model using MTD_T also overestimates the MTD, it is less likely to recommend overdose levels above $MTD + 1$ compared with TITE or conventional approach. CRM and hybrid designs have better convergence rates and recommend the target dose on MTD more often than EWOC when the true MTD is above the mid-level of the dose range considered. When the true MTD is below the mid-level, EWOC design is comparable to CRM or hybrid. EWOC controls the overdose proportion better than CRM or hybrid for all scenarios.

The effect of prior on the baseline hazard ratio (informative gamma prior vs. non-informative uniform prior) depends on the designs, MTD definition, true models and the length of the follow-up time. When the follow-up time is sufficient relative to the risk of time to toxicity, and the true MTD is above the mid-level of the dose range

considered (Model 1 or Model 2), the choice of prior on the baseline hazard rate has little effect on the performance of the CATE design using MTD_T for EWOC or hybrid; the performance is higher for CATE design using MTD_π for all designs or MTD_T for CRM when using informative prior than the non-informative prior. When the true MTD is below the mid-level of the dose range considered (model 3), the choice of prior on the baseline hazard rate has generally little effect on the performance of the CATE design using MTD_T or MTD_π, except in the case of EWOC using MTD_π when the performance is higher using informative prior than the non-informative prior. When the planned follow-up time is short, the choice of prior on the baseline hazard rate has little effect on the performance of the CATE design using MTD_π when the true MTD is below the mid-level. The performance of the CATE design using MTD_π is higher using informative prior than the non-informative prior when the true MTD is above the mid-level of the dose range considered.

Simulation shows that the CATE design based on the proportional hazards model using Weibull distribution under different true dose-toxicity models has generally high performance with better percentage of correct dose recommendation and low overdose proportion. Results are similar to those based on the proportional hazards model using exponential distribution. When data were generated using Weibull distribution but estimated using the exponential distribution (exponential four-parameter model), simulation shows that the CATE design maintains the generally high performance with better percentage of correct dose recommendation and low overdose proportion even under the inadequate model (exponential distribution). When data were generated using exponential distribution but estimated using the Weibull distribution, the CATE design maintains the generally high performance with better percentage of correct dose recommendation and low overdose proportion under this scenario.

The CATE design based on the accelerated failure time model using Weibull distribution under different true dose-toxicity models has similar performance to that based on the proportional hazards model using Weibull distribution. The CATE design has generally high performance with better percentage of correct dose recommendation and low overdose proportion. Designs using the MTD_T definition have better performance

than the corresponding one using the MTD_π. When the true MTD is above the mid-level of the dose range considered (model 1 and model 2), the hybrid design and CRM design have better convergence rates and recommend the target dose on MTD more often than EWOC. When the true MTD is below the mid-level of the dose range considered (model 3), EWOC design has comparable high performance as other designs. EWOC and the hybrid design generally provide a better safety protection in limiting higher dose for patients than the CRM designs do.

In conclusion, the proposed CATE has generally high percentage of correct dose recommendation and low overdose proportion. It significantly reduces the overall trial duration compared with the conventional approach. When the planned follow-up time is moderate or too short relative to the risk of time to toxicity, the CATE design (especially MTD_π for the short follow-up time) is superior to TITE or conventional approach, while TITE or conventional approach would overestimate MTD in certain situations.

8.2 Future Work

Based on the simulation results, the CATE design has good performance assuming that the DLT-free survival time for the susceptible patients follows a parametric distribution such as exponential or Weibull distribution. The CATE design assuming that the DLT-free survival time for the susceptible patients follows a semi-parametric distribution based on the proportional hazards model or accelerate failure time model could be studied.

The stopping rules for the phase I cancer clinical trials could be evaluated for the CATE design and see how the model performance are affected.

Bibliography

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

J. Babb, A. Rogatko, and S. Zacks. Cancer phase i clinical trials: Efficient dose escalation with overdose control. *Statistics in Medicine*, 17(10):1103–1120, 1998.

Y.K. Cheung and R. Chappell. Sequential designs for phase i clinical trials with late-onset toxicities. *Biometrics*, 56(4):1177–1182, 2000.

A. Mauguen, M.C. Le Deley, and S. Zohar. Dose-finding approach for dose escalation with overdose control considering incomplete observations. *Statistics in Medicine*, 30:1584–1594, 2011.

P. Chu, Y. Lin, and W. J. Shih. Unifying crm and ewoc designs for phase i cancer clinical trials. *Journal of Statistical Planning and Inference*, 139:1146–1163, 2009.

L. Edler. Statistical requirement of phase i studies. *Onkologie*, 13:90–95, 1990.

Y. Lin and W.J. Shih. Statistical properties of the traditional algorithm-based designs for phase i cancer clinical trials. *Biostatistics*, 2(2):203–215, 2001.

W.J. Shih and Y. Lin. Traditional and modified algorithm-based designs for phase i cancer clinical trials. In S. Chevret, editor, *Statistical Methods for Dose-Finding Experiments*. John Wiley and Sons, Ltd, 2006.

J.M. Heyd and B Carlin. Adaptive design improvements in the continual reassessment method for phase i studies. *Statistics in Medicine*, 18:1307–1321, 1990.

J. W. Boag. Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *Journal of the Royal Statistical Society*, 11:15–53, 1949.

J.W. Gamel, I.W. McLean, and S.H. Rosenberg. Proportion cured and mean log survival time as functions of tumor size. *Statistics in Medicine*, 9:999–1006, 1990.

K. Yamaguchi. Accelerated failure-time regression models with a regression model of surviving fraction: an application to the analysis of 'permanent employment' in japan. *Journal of the American Statistical Association*, 87(418):284–292, 1992.

J.P. Sy and J.M.G. Taylor. Estimation in a cox proportional hazards cure model. *Biometrics*, 56:227–236, 2000a.

Y Peng. Fitting semiparametric cure models. *Computational Statistics and Data Analysis*, 41:481–490, 2003.

Y. Wu, Y. Lin, S.E. Lu, C.S. Li, and W.J. Shih. Extension of a cox proportional hazards cure model when cure information is partially known. *Biostatistics*, 15(3):540–554, 2014.

J. O'Quigley and S. Chevret. Methods for dose finding studies in cancer clinical trials: A review and results of a monte carlo study. *Statistics in Medicine*, 10:1647–1664, 1991.

D. Faries. Practical modifications of the continual reassessment method for phase i cancer clinical trials. *Journal of Biopharmaceutical Statistics*, 20(17-18):2661–2681, 1994.

E.L. Korn, D. Midthune, T.T. Chen, L.V. Rubinstein, M.C. Christian, and R.M. Simon. A comparison of two phase i trial designs. *Statistics in Medicine*, 13:1799–1806, 1994.

S. Moller. 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. *Statistics in Medicine*, 15:1605–1618, 1995.

W.J. Shih, J.J. Lee, and Y. Lin. A hybrid of modified crm and ewoc approaches for phase i cancer trials. *International Biometrics Society ENAR Spring Meeting*, 1999.

J. Berkson and R. P. Gage. Survival curves for cancer patients following treatment. *Journal of the American Statistical Association*, 47:501–515, 1952.

A.Y.C. Kuk and C. Chen. A mixture model combining logistic regression with proportional hazards regression. *Biometrika*, 79:531–541, 1992.

J.P. Sy and J.M.G. Taylor. Estimation in a cox proportional hazards cure model. *Biometrics*, 56:227–236, 2000b.

Y. Peng, K.B.G. Dear, and J.W. Denham. A generalized f mixture model for cure rate estimation. *Statistics in Medicine*, 17:813–830, 1998.

C.S. Li and J.M.G. Taylor. A semi-parametric accelerated failure time cure model. *Statistics in Medicine*, 21:3235–3247, 2002.

D.R. Jones, R.L. Powles, D. Machin, and R.J. Sylvester. On estimating the proportion of cured patients in clinical studies. *Biometrie-Praximetrie*, 21:1–11, 1981.

A.I. Goldman. Survivorship analysis when cure is a possibility: A monte carlo study. *Statistics in Medicine*, 3:153–163, 1984.

M.E. Ghitany and R.A. Maller. Asymptotic results for exponential mixture models with long term survivors. *Statistics*, 23:321–336, 1992.

V.T. Farewell. The use of mixture models for the analysis of survival data with long-term survivors. *Biometrics*, 38(4):1041–1046, 1982.

N.H. Gordon. Maximum likelihood estimation for mixtures of two gompertz distributions when censoring occurs. *Communications in Statistics - Simulation and Computation*, 19:733–747, 1990a.

N.H. Gordon. Application of the theory of finite mixtures for the estimation of cure rates of treated cancer patients. *Statistics in Medicine*, 9:397–407, 1990b.