2017

On the designs of early phase oncology studies

BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

ON THE DESIGNS OF EARLY PHASE ONCOLOGY STUDIES

by

REVATHI NAYANTARA ANANTHAKRISHNAN

B.Sc., University of Pune, 1996
M.Sc., Indian Institute of Technology, 1998
Ph.D., University of Texas at Austin, 2003
M.A., Boston University, 2012

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2017

Approved by

First Reader

Michael LaValley, Ph.D.
Professor of Biostatistics

Second Reader

Mark Chang, Ph.D.
Adjunct Professor of Biostatistics

Third Reader

Gheorghe Doros, Ph.D.
Professor of Biostatistics

DEDICATION

I would like to dedicate this work to my nephew and niece, Kedhar and Sahana.

ACKNOWLEDGMENTS

Many things in life, it is said, are easier the second time around — I guess it is called experience. This has certainly been true of this PhD. I was no longer a naive, young graduate student in a completely new country with no idea of how the real world operated but was now someone well settled in the country with a full time professional career in the Pharmaceutical Industry with clear career goals and a clear vision and aspirations for my future. I knew why I wanted to do this PhD, what courses to take that were most relevant and useful to my work in oncology clinical trials, and most importantly how to choose the right adviser to help and support me in achieving my goals.

At the end of my post-doctoral fellowship in Biomathematics at the University of California at Davis, my future looked rather bleak to me. Getting an academic position in Physics in a good research oriented university and succeeding in it by continually obtaining funding, when a very low percentage of NIH/NSF grant applications were being funded, didn't seem like the most viable option for me. Neither did I want to apply to heavily teaching oriented university jobs and devote myself mainly to teaching with very little time for quality research. Getting a research position in Industry appeared to be an option to consider, but I had no idea what a theoretical physicist could do in Industry and how to go about obtaining a suitable position there. It was time for a lot of soul searching about what I liked, what I thought that I was good at and how I wanted to spend the rest of my life. The answer consistently

seemed to point towards Applied Mathematics, something that I had wanted to do after my MSc in Physics. My love of Mathematics was what had led me to do Theoretical Physics, and as a Theoretical Physicist, I strongly believed that I had the right background for Applied Math. Based on her knowledge of my biopolymer physics based modeling work on the shape, motility and division of cells with relevance to cancer, it was my sister who pointed out to me Biostatistics jobs in Industry as well as statistical classes in clinical trials being offered at the statistics department of UC Davis. I met and talked to the Professor, Prof Laurel Beckett, who was offering the Statistical Methods in Clinical Trials class, and then sat in on the entire course the quarter before my post-doc fellowship ended. She encouraged me considerably to go along this new path that I was exploring. Thus, began my long and interesting journey into Biostatistics.

At the end of my post-doc, I applied to many Pharmaceutical Industry jobs and finally obtained a position in the CRO Parexel. For the past 11 years, I have been working as a Research Biostatistician, almost 9 of them at the Oncology Business Unit in Pfizer and presently at Celgene, where I started 6 months ago after my move to NJ. My time at Pfizer gave me great and valuable exposure to the entire oncology clinical trials process from drug discovery and development and early clinical trials including pharmacological (PK/PD) modeling to late stage clinical trials and regulatory submissions. By working closely with clinicians and investigators/oncologists, other statisticians, SAS data analysts, PK/PD modelers, clinical data managers and many

others, I have been able to obtain a holistic view of the entire process from drug discovery to approval, and have also learnt a lot about the biology of solid tumors as well as lymphomas and leukemias, whose statistical endpoints can be extremely complex. I also worked in the field of Immunooncology in my last 2 years at Pfizer. The very careful, rigorous and high-level scrutiny of the many papers that I published as co-author at Pfizer has taught me a lot about the writing process and has enabled me to improve my technical writing considerably and to produce high quality, high impact scientific papers. The most gratifying part of my time/work at Pfizer was, however, that 2 drugs that I worked on as part of the statistical team, Bosutinib for CML and Inotuzumab for ALL, were submitted to the US FDA for drug approval. My initial disappointment right after the post-doc in realizing that an academic career in Biophysics may not work out for me has been more than compensated for by knowing that as part of these teams, I have directly been able to impact some lives with my present work. Sometimes, the future that unfolds turns out to be infinitely richer and more satisfying than what one initially planned or imagined for one's future, and I am very grateful for that.

As soon as I started my job in the Pharmaceutical Industry, I also began taking classes at Boston University (BU) as a part-time student. I enrolled in 2010 for my MA in Biostatistics, completed it at the end of 2011 and enrolled in 2012 as a part-time PhD student. The synergy of using what I learnt at school for my work and using my work experience to do well at school has been one of the most rewarding experiences of

this journey. In late 2014, after the completion of a total of ~18 graduate courses, including courses in Mathematics, Biostatistics and Epidemiology, and 2 Applied and 2 Theory qualifying exams, I started my dissertation work with Prof. Michael LaValley. I feel very privileged to have had such an outstanding mentor. He is one of the most patient people that I have worked with, and discussions with him were always very useful and insightful. Most of the credit for my PhD thesis experience being so smooth and fulfilling goes to him, and I thank him immensely for his support and guidance. I also thank my committee members Profs. Gheorghe Doros, Joe Massaro, Mark Chang and Dr. Stephanie Green for all their valuable advice and support. I really enjoyed interacting with them and appreciated their input tremendously. I also thank Prof. Howard Cabral for his great support and encouragement through the years. He has often advised me to be patient and kind with myself and to keep proceeding with my work and studies. I would like to thank my course adviser Prof Gagnon for his support during my course work stage. The BU Biostatistics department has an excellent faculty and staff and I have really enjoyed my time at BU. I would like to thank MyHanh, Marisa and Kaitlyn for all their help through these years at BU. I also thank my many classmates over the years at BU. Some of whom I have shared homework, project and study sessions and really enjoyed the interactions include Avery, Danielle, Greg, Hanna, Jing, Joseph, Katya, Mani, Michael, Pranab, Siyan, Soudeh and Susan. Being surrounded by young, inquiring minds has definitely been very invigorating and energizing and has always made me feel much younger than what I am.

I owe a huge debt of gratitude to one of my managers at Pfizer, Stephanie Green. She is not only a highly competent and sharp statistician, whose co-authored book on the Statistics of Oncology Clinical Trials is a standard in the field, but also an excellent mentor. She very skillfully guided me for many years through the maze that is oncology clinical trials, from the statistics of solid tumors to that of liquid tumors (lymphomas and leukemias). I have learnt a lot from her by interaction and osmosis! She has also unstintingly encouraged and supported my PhD work and I have had numerous useful discussions with her about this work. I also thank another manager at Pfizer, Jane Liang White. She is one of the nicest people I have worked with and an extremely competent statistician; working with her was a great pleasure. She greatly supported hiring me as a statistician in the Pfizer Immunooncology group and I am grateful that she still cares about my career and progress. I thank Dmitri Pavlov and Ke Zhang, with whom I worked closely on the Immunooncology projects at Pfizer, for their support. I thank the present VP of Oncology Statistics at Pfizer, Enayet Talukder, as well as the previous VP, Janis Grechko, who is also my present VP at Celgene, for encouraging my doing the PhD on the side. I thank my manager at Celgene, Dale Song, and his manager Tommy Fu, for all their support and encouragement. I thank my manager at Inventiv Health, James Trammel, for all his support and guidance of my career and his valuable insights while I was with Inventiv. It was a pleasure interacting with him. I owe a huge debt of gratitude to Andrew Strahs, my manager at Wyeth Research, who took a huge chance on me when I started out in the Pharmaceutical Industry and told him at the interview that I knew some basic

oncology, some mathematical modeling but barely any statistics. He told me that he wasn't looking for just a statistician but a scientist who was willing to learn and grow. His broad thinking and the fact that he didn't box me based on my degree impressed me a huge deal and I have been in his debt ever since. I also thank my manager at Parexel and colleague at Pfizer and friend for many years now, Eric Leip, who has mentored me with regard to work and school in uncountable ways. He has been very patient with my many questions and has helped me considerably. I also thank all my present and past colleagues in all my project teams at work for their support and help throughout these years. There are certainly many colleagues from whom I have learnt a lot and I am grateful for the stimulating interactions with all of them.

I thank all my friends who have continued to support me — I apologize in advance for not mentioning them individually but they are too numerous to mention here! It has meant a great deal to me and I am indebted to them all. I thank my close friends from my graduate school at UT Austin and my post-doc at UC Davis who have continued to encourage me in my further pursuits. Although some of my friends thought that I was completely crazy to do another PhD in the hard Sciences after a rigorous PhD in Physics from the Center for Nonlinear Dynamics at UT Austin and wondered why I couldn't learn the needed math and statistics at work or by myself, they nevertheless humored me and supported me fully in my endeavors. I have not doubted that I made the right decision however. Soon after I joined the Pharmaceutical Industry, it became clear to me that it could be difficult to progress professionally as a Biostatistician in

the Pharmaceutical Industry without having a PhD in Biostatistics and realized that it would be a good idea to go ahead and obtain one while I had the opportunity to do so. In this process, however, I banked extremely heavily on my strong mathematical background and mathematical intuition developed from my years doing Theoretical Physics to get me through the challenging Mathematical Statistics PhD core courses such as Estimation Theory and Hypothesis Testing and the corresponding PhD Theory qualifying exam as well as the PhD in general while working full time.

I also thank my previous Professors, advisers, mentors and colleagues in Physics who have supported me through the years and who have helped me build a strong mathematical foundation and the confidence that with the kind of broad background, training, problem solving abilities and other tools and transferable skills that a graduate Physics education provided, I could branch out into any field that I wanted to pursue.

I also thank all the doctors and specialists at Mass General (MGH), Spaulding Rehab, Harvard Vanguard and the Summit Medical Group for monitoring my health very closely and hence allowing me to lead as normal a life as is possible.

Finally, I thank the most important people in my life — my family. I thank my dad for pushing, supporting, coaxing and cajoling me into doing Science all my life, for strongly believing in my scientific capabilities and for exhorting and inspiring me to

do better than my very best in all my professional and personal undertakings. Importantly, I thank him for passing on three of his traits to me either via his genes or via his actions, or likely a combination of both: an unquenchable thirst for pursuing knowledge and attaining wisdom, a deep love of Science, and an infinite capacity for sheer hard work and perseverance. Over the past many years, he has strongly instilled in my brain the idea that my 'potential intellect' aka my inherent talent/capability/intelligence/creativity (akin to potential energy in Physics) did not matter as much as how much of all this I was actually able to convert into 'kinetic intellect' aka useful/productive work (akin to kinetic energy in Physics) via hard work and perseverance. I have more recently read articles on the fixed mindset versus the growth mindset and realize that my father has been advocating the growth mindset to me for a very long time. My mom is smart, sensible, practical and hardworking and is the rock that anchors me in good times and in bad times. She is the person I talk to when I have any joys to share or difficulties to unburden. Her support and love are strong and unwavering and they are the immutable laws of my universe that I heavily rely upon. I simply couldn't do without my sister. She is smart, sensible, practical and hardworking. She has always been and will always be my best friend and talking to or spending time with her and her family is always the best part of my day. Her support and love are constant and steady and the number of times during this PhD that she has reminded me that when the going gets tough, the tough get going is uncountable. I thank my brother-in-law for all his support, encouragement and discussions throughout these years, which are all very much

appreciated. Lastly but most importantly, I thank Kedhar and Sahana, who bring immeasurable joy and meaning to my life. I am very grateful to be a part of their lives, very proud of what they have accomplished so far, and very excited to watch their growth, progress and accomplishments in the coming years. I truly hope that all the hard work, persistence, dedication and passion that have gone into this degree and this thesis will inspire and propel them to really great heights in whatever they choose to pursue. To them, I dedicate this thesis and to them I wish an unimaginably bright and beautiful future.

ON THE DESIGNS OF EARLY PHASE ONCOLOGY STUDIES

REVATHI NAYANTARA ANANTHAKRISHNAN

Boston University Graduate School of Arts and Sciences, 2017

Major Professor: Michael LaValley, Professor of Biostatistics

ABSTRACT

This thesis focuses on the design, statistical operating characteristics and interpretation of early phase oncology clinical trials. Anti-cancer drugs are generally highly toxic and it is imperative to deliver a dose to the patient that is low enough to be safe but high enough to produce a clinically meaningful response. Thus, a study of dose limiting toxicities (DLTs) and a determination of the maximum tolerated dose (MTD) of a drug that can be used in later phase trials is the focus of most Phase I oncology trials. We first comprehensively compare the statistical operating characteristics of various early phase oncology designs, finding that all the designs examined select the MTD more accurately when there is a clear separation between the true DLT rate at the MTD and the rates at the dose levels immediately above and below. Among the rule-based designs studied, we found that the 3+3 design under-doses a large percentage of patients and is not accurate in selecting the MTD for all the cases considered. The 5+5 a design picks the MTD as accurately as the model based designs for the true DLT rates generated using the chosen log-logistic and linear dose-toxicity curves, but requires enrolling a larger number of patients. The model based designs examined, mTPI, TEQR, BOIN, CRM and EWOC designs, perform well on the whole, assign the maximum percentage of patients to the MTD, and pick

the MTD fairly accurately. However, the limited sample size of these Phase I oncology trials makes it difficult to accurately predict the MTD. Hence, we next study the effect of sample size and cohort size on the accuracy of dose selection in early phase oncology designs, finding that an adequate sample size is crucial. We then propose some integrated Phase 1/2 oncology designs, namely the 20+20 accelerated titration design and extensions of the mTPI and TEQR designs, that consider both toxicity and efficacy in dose selection, utilizing a larger sample size. We demonstrate that these designs provide an improvement over the existing early phase designs.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

Introduction

This statistical thesis focuses on the design, statistical operating characteristics and interpretation of early phase oncology clinical trials. To put our work in perspective, we begin by explaining the genetic and cytoskeletal changes that take place in cancer cells and that result in their hallmarks, including the key one of limitless cell division. We then describe the process of cell division. Cell division is a very precise and controlled process, and the changes in its control mechanisms, such as in specific genes and proteins, that lead to abnormal cell division in cancer cells can be utilized in the development of anti-cancer drugs. Subsequently, we discuss anti-cancer drugs as well as their rigorous process of testing i.e. oncology clinical trials. Finally, we discuss the aims and scope of this thesis. In brief, the aims of this thesis include gaining a comprehensive understanding of the statistical operating characteristics of several existing early phase oncology designs, systematically studying the effect of cohort and sample size on the accuracy of dose selection in some existing designs, and then proposing new integrated Phase 1/2 oncology designs that are improvements over the existing designs in various ways. This thesis is unique in its comprehensive survey of existing early phase oncology designs and in its proposal of new designs that are based on a thorough statistical understanding of the existing designs.

Cancer[1]

Cancer has long been a dreaded disease and has long been synonymous with a death sentence. However, a tremendous amount of progress has been made in the past 50 years and today at the beginning of the 21st century, this is no longer always the case. Today, there are many patients diagnosed with cancer who proceed to live their full functional life spans. For example, the drug Gleevec has turned chronic myeloid leukemia (CML) into a chronic, manageable disease for many CML patients [1]. The prognosis of the cancer diagnosis usually depends on the stage[2] at which the cancer is detected. In many cases, early detection followed by immediate treatment results in the best outcome [5, 6, 7, 8, 9]. Early detection tools such as a mammogram for breast cancer, Pap test for cervical cancer, PSA test for prostate cancer, flexible sigmoidoscopy and colonoscopy for colorectal cancer aim to help improve the outcome from these cancers [10]. The typical treatment regimen for cancer consists of anti-cancer drugs, such as chemotherapy, hormonal therapy, targeted therapy, and or radiation and or surgery. There are also other new innovative therapies currently being investigated, which we discuss later. Known risk factors for cancer include cigarette smoking and tobacco use, exposure to radiation, inhalation of asbestos fibers, infections and genetics for e.g. the presence of mutations in the BCRA1 and

---

[1] This section and the following sections until the section"Aims and Scope of this Thesis" are based on my two review articles on oncology trials [23] and [50].

[2] For most solid tumors, the TNM (tumor, node, metastasis) staging is used, which depends on the tumor size, how many lymph nodes are affected and if the tumor has metastasized [2]. The results from the TNM staging are combined to determine the final staging result of Stage IA, IB, IIA etc. and this conversion differs for different solid tumor types. For liquid tumors, the staging criteria and their definitions are different for different leukemias [3] and lymphoma [4].

BCRA2 genes significantly increases the risk of breast cancer [11]. Other risk factors for cancer can include diet, drinking, obesity, exercise and environmental factors.

So, what is cancer? Firstly, cancer is not a single disease but a complex family of diseases that are primarily characterized by uncontrolled cell division. The molecular underpinnings and mysteries of cancer are still being unraveled. However, all the research done so far generally points to cancer being a disease of genetic alterations such as chromosomal abnormalities (e.g. genetic mutations) [12]. These mutations lead to abnormal cell growth and division and to nearby tissue invasion by these cells. Some of these cells may break away from the original tumor and travel to far off sites, which is called metastasis. The number and type of mutations can vary greatly even between two patients diagnosed with the same cancer e.g. lung cancer. In addition, the same patient may have different mutations at different time points after the initial diagnosis, as the mutations can evolve over time. This is why patients often develop resistance to the targeted therapy that they are being treated with — the specific mutation that the drug was initially targeting has likely changed over time and the patient needs to then be treated with another drug that targets the new mutation. Thus, the key to treating cancer seems to lie in identifying the unique genetic signature of each patient's cancer at the given time point and in personalizing the treatment plan accordingly.

Hallmarks of Cancer and Genetic and Cytoskeletal Changes in Cancer Cells

The main hallmarks of cancer cells are thought to include self-sufficiency in growth signals, insensitivity to anti-growth signals, tissue invasion and metastasis, limitless replicative potential, sustained angiogenesis (formation of new blood vessels) and evasion of apoptosis (cell death) [13]. These hallmarks have been revised in a more recent paper [14] by the same authors as those of the original paper to include abnormal metabolic pathways and evading the body's immune system (Figures 1 and 2 show the original hallmarks and the emerging hallmarks respectively, taken from [14]).

Figure 1 Original Hallmarks of Cancer (taken from Hanahan 2011 [14])

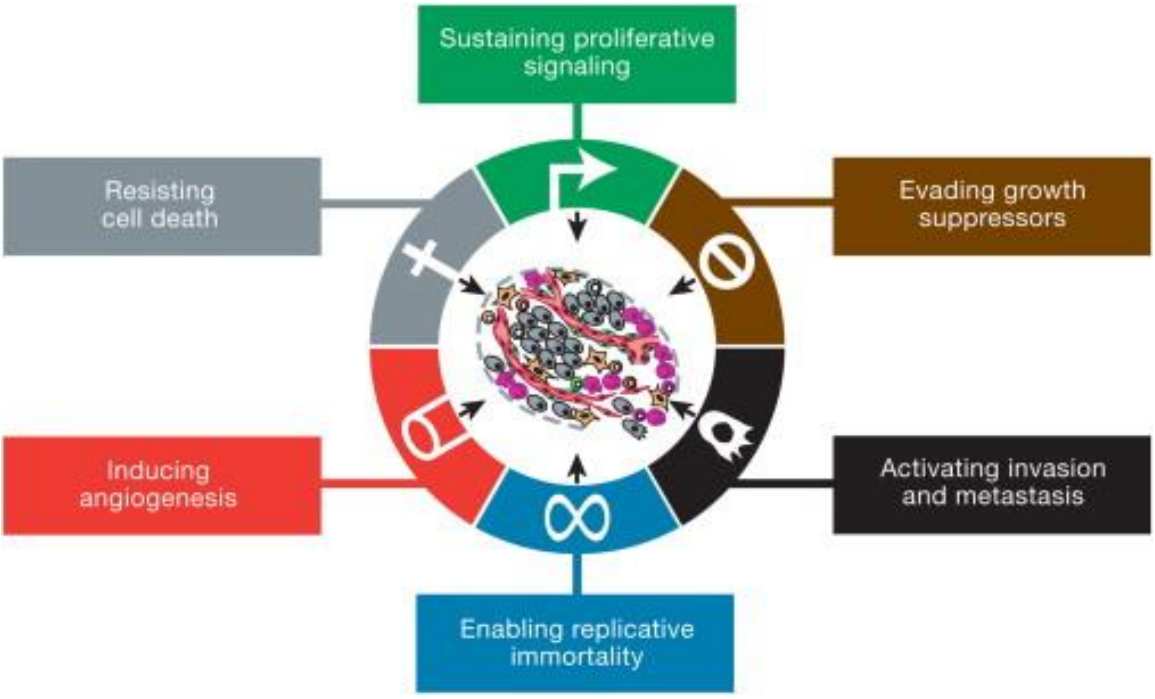
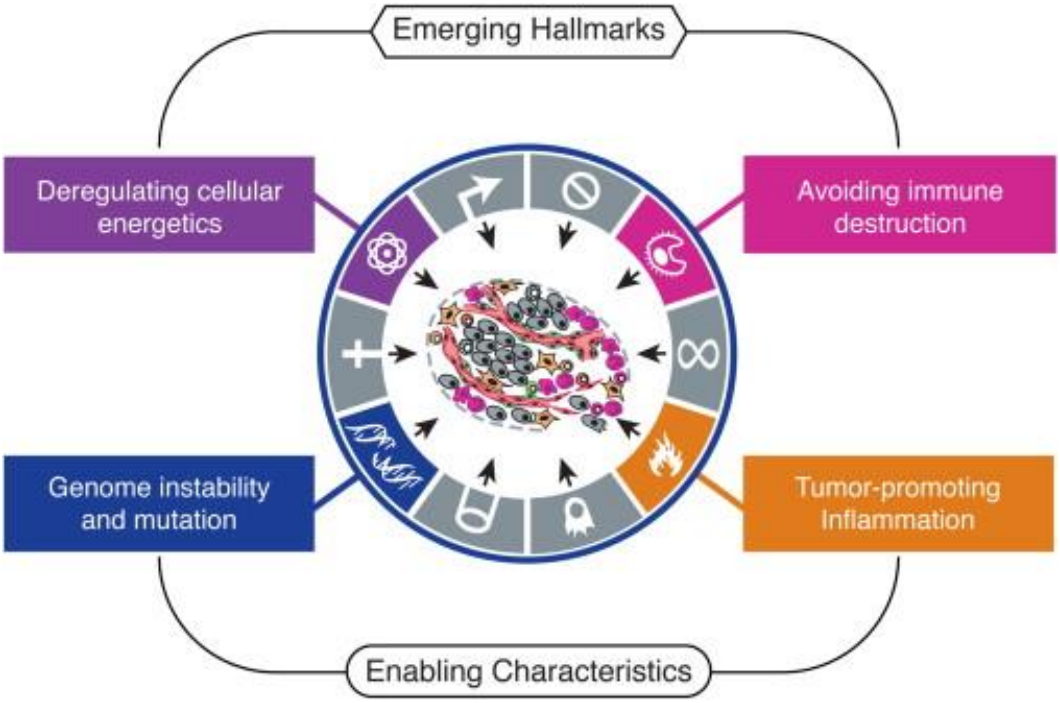


Figure 2 Emerging Hallmarks of Cancer (taken from Hanahan 2011 [14])



Cancer cells exhibit these hallmarks due to the genetic and cytoskeletal changes that take place in them compared to their normal counterparts. Most of the cell's genetic material is contained in the genes that make up the chromosomes. The chromosomes are found in the cell's nucleus, a dense body in the cell interior. The main genetic changes in cells that contribute to the development of cancer relate to oncogenes, tumor suppressor genes, DNA repair genes and apoptosis genes, which are genes that control cell death [15, 16]. Proto-oncogenes and tumor suppressor genes are critical in regulating cell division, and mutations in these genes can render them unable to function normally. In general, when proto-oncogenes are activated due to gain of function mutations, they become oncogenes, while when tumor suppressor genes are inactivated due to loss of function mutations, they can lead to cancer. The functions

of various proto-oncogenes include regulating cell growth and apoptosis, and controlling the timing and frequency of cell division via signals sent to the nucleus; when a proto-oncogene becomes an oncogene, it causes the signaling mechanism that instructs the cell when to divide to function abnormally, resulting in excessive cell division. The presence of a single oncogene in a cell is not likely to make it cancerous but promotes cell division, which may aid one or more daughter cells down the line in acquiring additional mutations and becoming cancerous. In general, tumor suppressor genes function to stop cell division when it is no longer necessary. Hence, when a tumor suppressor gene is mutated and does not function normally, it can lead to abnormal cell division. Thus, a cancer cell with oncogenes and mutated tumor suppressor genes is often compared to a car whose accelerator pedal is permanently pushed down and whose brakes do not function. Examples of oncogenes include the BCR-ABL oncogene, which promotes cell growth and division through tyrosine kinase activity and is associated with CML, and the HER2 oncogene, which promotes cell growth and division by over expression of signaling kinase due to gene amplification and is usually associated with breast, cervical, endometrial and ovarian cancer. Examples of tumor suppressor genes include the BRCA1, BRCA2 genes, whose mutations are associated with breast and ovarian cancer, and the p53 gene, whose mutations are associated with several cancers such as breast, colorectal, liver, lung and ovarian cancers. Genes that are involved in DNA replication and repair are also vital; when these genes are mutated, their ability to automatically correct mutations that may cause cancer is lost. Apoptosis genes are also very important with regard to

cancer because some cancer cells are able to divide forever. Two main genes whose mutations can lead to the loss of apoptosis are p53 and BCL-2. Some cancer cells are able to divide indefinitely by activating the enzyme telomerase, which controls the length of the telomeres – telomeres act like caps at the ends of chromosomes [17]. When telomerase is activated, it prevents the shortening of the telomeres and enables these cancer cells to keep on dividing; in normal cells, when the telomeres shorten to or beyond a certain threshold, further cell division does not take place.

Cytoskeletal changes also occur in cancer cells [18]. The cytoskeleton is a dynamic, composite polymeric network spanning the interior of eukaryotic cells, which are cells that have a nucleus. This polymeric network acts as a skeleton to provide the cell structure and shape. The biopolymers actin, microtubules and intermediate filaments comprise the cytoskeletal network [18]. These cytoskeletal polymers function synergistically to carry out several key cellular activities such as motility, division and intracellular transport [18]. The actin cortex, a shell-like dynamically crosslinked polymer network situated just below the cell membrane, is made out of semiflexible actin filaments. Actin filaments can also form stiff bundle like structures known as stress fibers. The actin cytoskeleton, comprising the actin cortex and stress fibers, is very important in cell elasticity and shape, cell motility and muscle contraction [19]. Microtubules are rigid rod like structures that start at the centrosome, located near the nucleus, and extend to the cell membrane in a radial array-like configuration. They are key in cell division, by enabling the movement and separation of sister

chromatids to opposite sides of the cell by attachment proteins called kinetochores. They also act as a set of tracks for organelles and vesicles to move on, and hence are vital in intracellular transport. The flexible intermediate filaments (IFs), which are intermediate in size (diameter) between actin filaments and microtubules, comprise a class of filaments such as vimentin, neurofilaments, nestin, keratin and desmin; different cells can have different IFs. IFs provide mechanical support to the cell membrane when it comes into contact with other cells. IFs have high tensile strength and long-range elasticity and are crucial when the cell undergoes large deformations. Frequently, the actin cytoskeleton is remodeled in many cancer cells, with the actin cortex beneath the cell membrane usually being less extensive in the cancer cells than in their normal counterparts. This results in the altered elastic/viscoelastic and other properties of these cancer cells, and consequently in their altered functioning [20, 21, 22, 23]. Since the actin network and its dynamics are crucial in cell shape and movement, these altered properties affect cell shape, movement and growth. The smaller shear modulus and the more liquid-like viscoelastic properties rather than solid-like elastic properties of these cancer cells compared to their normal counterparts allow them to squeeze through gaps much more easily and move more quickly than their normal counterparts [24]. Some cancer cell types also divide more quickly than their normal counterparts. In addition, changes in microtubule stability and disruption of processes that microtubules are involved in have been observed in many cancers [25, 26]. Further, the roles for IFs in cancer include the following: vimentin plays a role in lung cancer, keratin plays a role in apoptosis and nestin plays

a role in cancer cell migration [27, 28, 29, 30].

In summary, the genetic and cytoskeletal changes that occur in cancer cells result in their altered functioning, including cell division.

Cell Division

Cell division is usually initiated when growth factors attach to specific receptors on the cell surface and send signals to the cell's nucleus, leading to a series of events that culminate in the creation of two cells with identical genetic material [31]. In eukaryotic cells, the cell cycle, the process by which one cell divides into two, can be divided into various stages [32]. The cell first prepares for division in interphase. The stages of interphase are G1 (Gap 1 stage, where there is growth and preparation of chromosomes for replication), S ("Synthesis", where DNA replication occurs) and G2 (Gap 2 stage, where the cell prepares for mitosis). The next stage is the M stage or "Mitosis" stage, where the nuclear chromosomes separate. Mitosis is split into four main stages: prophase, metaphase, anaphase and telophase. Over the course of these four stages, the centrosomes are first separated and migrate to opposite ends of the nucleus, the chromosomes are then aligned at the cell center and finally the sister chromatids are separated and pulled to opposite sides of the cell by kinetochores. The end result of mitosis, where microtubule dynamics play an essential role, is that duplicate copies of chromosomes are split equally. During the next stage called cytokinesis, the cell's cytoplasm gets divided; this finally results in the creation of two genetically identical daughter cells. In some cells, there can also be a post-mitotic G0

phase where the cell has left the cell cycle and is resting and not dividing. The cell cycle is a very complex and regulated process, whose precise orchestration involves numerous checkpoint proteins such as cyclins, cyclin dependent kinases as well as genes such as proto-oncogenes and tumor suppressor genes. Hence it is not surprising that mutations in the proto-oncogenes, tumor-suppressor genes or cell cycle checkpoints that are key in controlling this accurate process can lead to abnormal cell division. To inhibit abnormal cell division in cancer cells, the different anti-cancer drugs that have been developed may target different parts or stages of the cell cycle, as discussed in the next section. The drugs can also make use of the hallmarks or special characteristics exhibited by cancer cells in order to inhibit tumor growth, for e.g. targeting specific genes (e.g. oncogenes) and proteins that are involved in signaling the cancer cell to continually divide, inhibiting angiogenesis around the tumor to slow or halt the tumor growth, or stimulating the body's suppressed immune system to attack cancer cells.

Oncology Drugs

The initial drugs that were developed and used to treat many different cancers were mainly chemotherapies [33]. Chemotherapy can be given either before or after surgery or radiation therapy. When it is given before surgery to shrink the primary tumor, this is called neo-adjuvant therapy. When it is given after surgery or radiation therapy to eradicate any cancer cells that may still remain in the body and to minimize

recurrence, this is called adjuvant therapy. When chemotherapy is given after the cancer has metastasized, this is called chemotherapy in the metastatic setting. When chemotherapy or another anti-cancer drug is the first drug given to the patient after diagnosis and is the primary or best therapy for the specified cancer, it is called 1st line therapy. When chemotherapy or another anti-cancer drug is given to a patient after he or she has progressed on the first anti-cancer drug, it is called 2nd line therapy and so on. Chemotherapies are not very cell specific and can target both normal and cancer cells – they usually target any actively or quickly dividing cells such as cancer cells and normal hair, nail and gut cells. The destruction of gut cells in the lining of the digestive system is what causes vomiting in many patients on chemotherapy. In general, chemotherapies work by targeting and damaging the DNA, RNA or microtubules in cells; different chemotherapy drugs can target different phases of the cell cycle. The main classes of chemotherapies include alkylating agents, anthracyclines, anti-metabolites, anti-microtubule agents and topoisomerase inhibitors [34]. Alkylating agents damage DNA, and this stops the cancer cells from dividing. They work in all phases of the cell cycle. As an example of an alkylating agent, cyclophosphamide forms DNA crosslinks; this interferes with DNA replication, and leads to cell death. Anthracyclines are thought to work in multiple ways to inhibit cancer [35]. They can inhibit DNA or RNA synthesis by inserting between base pairs in the DNA/RNA strand. They can damage DNA by creating free oxygen radicals. They can prevent DNA transcription and replication by inhibiting topoisomerase II. They can also stop the response to DNA damage by promoting histone eviction from

chromatin. They work in all phases of the cell cycle. As an example of an anthracycline, daunorubicin can inhibit DNA synthesis by inserting between base pairs in the DNA strand. It can also prevent DNA transcription and replication by inhibiting topoisomerase II. Finally, it can stop the response to DNA damage by promoting histone eviction. Anti-metabolites interfere with DNA and RNA synthesis and hence stop cell division. They act during the S phase of the cell cycle. As an example of an anti-metabolite, azathioprine inhibits an enzyme necessary for DNA synthesis and stops cancer cells from dividing. Mitotic inhibitors, as their name implies, disrupt mitosis in the M phase [36]. Most of these mitotic inhibitors disrupt microtubules and microtubule dynamics, given their crucial role in mitosis. Taxanes and plant alkaloids are examples of mitotic inhibitors. As an example of a taxane, paclitaxel binds to tubulin, the major building block of microtubules, stabilizing microtubules and preventing their disassembly. Due to this, the chromosomes are unable to form a spindle in metaphase. Thus, mitosis is arrested and cancer cells are stopped from dividing. As an example of a plant alkaloid, vincristine binds to tubulin, preventing microtubule disassembly. This arrests mitosis in metaphase and stops cell division. Topoisomerase inhibitors are drugs that disrupt the activity of the topoisomerase enzymes, namely topoisomerase I and topoisomerase II, which help separate and unwind DNA strands during transcription or replication. They act during the S phase and G2 phase of the cell cycle. Examples of topoisomerase I inhibitors include irinotecan and topotecan, while examples of topoisomerase II inhibitors include doxorubicin and aclarubicin. For example, irinotecan inhibits topoisomerase I, which

results in the inhibition of both DNA transcription and replication. One of the ways doxorubicin is thought to work is by inhibiting topoisomerase II; this results in blocking DNA transcription and replication.

The newer targeted therapies are more specific and usually inhibit cell proliferation rather than kill cancer cells [37]. The National Cancer Institute's (NCI) dictionary defines a targeted therapy as: "A type of treatment that uses drugs or other substances to identify and attack specific types of cancer cells with less harm to normal cells. Some targeted therapies block the action of certain enzymes, proteins, or other molecules involved in the growth and spread of cancer cells. Other types of targeted therapies help the immune system kill cancer cells or deliver toxic substances directly to cancer cells and kill them. Targeted therapy may have fewer side effects than other types of cancer treatment. Most targeted therapies are either small molecule drugs or monoclonal antibodies." Early examples of targeted therapies are trastuzumab (Herceptin, a monoclonal antibody) for HER2+ breast cancer [38] and imatinib (Gleevec, a small molecule drug) for CML [39]. In HER2+ breast cancer patients, the HER2 gene is overexpressed in breast cancer cells, which means that too many copies of the gene are made. This gene encodes for a protein called the HER2 receptor. Due to the presence of too many HER2 receptors on the cancer cell's surface, the signals being sent to the nucleus that direct the cell to divide are greatly amplified. This leads to abnormal and limitless cell division. Herceptin is a large molecule monoclonal antibody and cannot enter the cell. It attaches itself to

the HER2 receptors on the cell surface of breast cancer cells and intercepts these signals that are being sent to the cell's nucleus and that are causing the cell to keep on dividing. Thus, by blocking these signals, Herceptin can slow or halt the cancer growth. In most patients with CML, the ABL and BCR genes are fused, forming an oncogene called BCR-ABL. This gene encodes for the BCR-ABL protein, which is a constitutively active tyrosine kinase. This means that the tyrosine kinase is continuously active and promotes unregulated cell division and the initiation of cancer. Gleevec is a small molecule tyrosine kinase inhibitor and can enter the cell. It binds close to the ATP binding site of BCR-ABL and can decrease the BCR-ABL enzyme activity and hence inhibit the uncontrolled proliferation of cancer cells. As another example of a targeted therapy, bevacizumab (Avastin) is a monoclonal antibody that is used mainly in the treatment of colorectal cancer and lung cancer, and is an angiogenesis inhibitor. It slows cancer growth by binding to a protein often overexpressed by tumor cells called vascular endothelial growth factor (VEGF) that promotes the growth of new blood vessels around the tumor [40].

An antibody-drug conjugate consists of a cytotoxic agent attached to an antibody, with trastuzumab emtansine and brentuximab vedotin being examples [41]. In trastuzumab emtansine, the cytotoxic agent DM1 is attached to trastuzumab. Trastuzumab targets HER2+ cancer cells and so the cytotoxin DM1 is also delivered only to these cancer cells. DM1 then enters the cells, binds to tubulin and stops cell division [42]. In brentuximab vedotin, the cytotoxic agent monomethyl auristatin E

(MMAE) is attached to the antibody cAC10. cAC10 binds to CD30, which is commonly found on the surface of cancer cells. MMAE is then released into the tumor cell, where it blocks tubulin polymerization and stops cell division.

Hormonal therapy can be effective in treating cancers that may be hormonally driven such as breast, prostate and endometrial cancer [43, 44]. These hormone therapies commonly block or remove the hormones that encourage cancer growth. Selective Estrogen Receptive Modulators (SERMs) such as tamoxifen to treat estrogen receptor positive (ER+) breast cancer and anti-androgens such as flutamide and bicalutamide to treat prostate cancer are examples of hormonal therapies. Tamoxifen binds to the estrogen receptor and blocks the action of estrogen, thus inhibiting cancer growth that is fueled by the action of estrogen, such as in ER+ breast cancer. The anti-androgens flutamide and bicalutamide bind to the androgen receptor and block the action of androgen, thus inhibiting prostate cancer growth that is fueled by the action of androgen. Aromatase inhibitors such as letrozole and anastrozole for the treatment of ER+ breast cancer are also examples of hormonal therapies. These drugs bind to the aromatase enzyme and suppress the conversion of androgens to estrogens by the aromatase enzyme; this results in slowing or stopping cancer growth in ER+ breast cancer patients.

The field of cancer immunotherapy uses the body's own immune system to attack the cancer cells. In most cancer patients, the immune system is severely compromised

and does not function normally because of the presence of an immuno-suppressive environment around cancer cells. Immunotherapy can work, in general, by either stimulating the body's immune system to help fight cancer or providing the immune system with antibodies to help fight cancer. With regard to re-activating the body's immune system and the development of immunotherapies, a crucial role is currently being played by the programmed death 1 (PD-1) receptor protein, that is expressed on the surface of activated T cells, and its ligands, programmed death ligand 1 (PD-L1) and programmed death ligand 2 (PD-L2). T cells are lymphocytes that are active participants in the body's immune response and can recognize and attack cancer cells. Many cancer cells produce excess PD-L1 and or PD-L2 which can bind to PD-1, inactivating the T cells and preventing them from attacking the tumor. Nivolumab, for example, is an anti-PD-1 monoclonal antibody and immunomodulator approved for the treatment of melanoma and squamous non-small cell lung cancer that prohibits PD-L1 and PD-L2 from binding to the PD-1 receptor on T cells [45]. This enables the T cells to be activated and attack the tumor.

Immunotherapy generally consists of three main groups: cell-based therapies (e.g. cancer vaccines), antibody therapies and cytokine therapies [46, 47, 48]. An example of a cancer vaccine is Provenge, which is used to treat prostate cancer and which stimulates the body's immune system to attack the prostate cancer cells. Some antibody therapies such as Herceptin are examples of immunotherapies. Monoclonal antibodies, which are man-made versions of immune proteins, can work as

immunotherapies in different ways as follows: by a) attaching to cancer cells and blocking the signaling pathway that causes abnormal growth and division e.g. Herceptin b) attaching to cancer cells and then attracting immune cells to attack these cancer cells e.g. alemtuzumab or c) provoking the immune system via the inhibition of immune system checkpoints e.g. nivolumab, ipilimumab [49]. Examples of cytokines are Interleukins and Interferons such as interleukin-2 (IL-2) and interferon-α (IFN-α). Since these cytokines modulate the immune response and can aid anti-tumor activity, these cytokines can be used as drugs to stimulate an immune response and treat cancer e.g. use of high-dose IL-2 therapy in the treatment of renal cell carcinoma and melanoma.

Oncology Clinical Trials and Cancer Treatments

Before a new anti-cancer drug can be brought to the market, it must undergo rigorous clinical testing in humans via clinical trials to evaluate its safety and efficacy [50, 51] (Figures 3 and 4). Traditional clinical trials in humans typically consist of Phase I, II and III trials, while the newer adaptive designs are more flexible and do not always have three phases, such as a seamless Phase I/II trial where phases I and II are combined or a seamless Phase II/III trial where phases II and III are combined. Patients must provide written informed consent to participate in these clinical trials and strict eligibility criteria are followed for patient selection. However, before testing the drug in humans, in vitro studies as well as animal studies of the drug are

carried out to obtain a preliminary understanding of its adverse effects and to help determine an initial safe dose of the drug in humans. These initial pre-clinical studies may also aim to understand the pharmacokinetic (PK) and pharmacodynamic (PD) properties of the drug i.e. the effect of the body on the drug and the effect of the drug on the body respectively. A Phase I study or trial, conducted on a small group of say 15–30 patients, focuses on drug safety (adverse events of the drug), drug dosing (what dose and route of administration e.g. oral or IV may be the best for the study drug), as well as the PK/PD properties of the drug in humans [52]. Sometimes biomarker studies may also be conducted as part of a Phase I study. The efficacy of the drug may be observed in Phase I trials but is not traditionally their focus. Based on the safety properties of the drug and the results of the Phase I trial, the drug is then tested in a Phase II setting. A Phase II trial typically enrolls 100s of patients and focuses on observing trends in the safety and efficacy of the drug. If the drug shows acceptable safety properties and shows positive trends in efficacy, the drug is tested in a Phase III setting. A Phase III trial is large and can enroll 1000s of patients. It is typically a randomized parallel-arm superiority trial, where patients are randomized to the standard of care (SOC) in one arm and to the new study drug in the other arm, and the aim is to demonstrate the superiority of the study drug to the SOC. In most cases, obtaining market approval of the study drug entails showing regulators the results of such a Phase III randomized clinical trial (RCT), since an RCT is the gold standard for clinical testing. Based on the results of the Phase III RCT, the sponsor company can submit the drug to the regulatory agency for approval. The regulatory

agency, using a panel of experts, carefully analyzes the risk to benefit ratio of the drug and decides accordingly. Once the cancer drug is approved in a country/region, it becomes available to all the patients there who require it.

Personalized cancer therapy seems to be where the future of cancer treatment lies. This implies that the treatment regimen, which may include targeted therapies, immunotherapies, cell based therapies (for e.g. injecting T cells into the body to fight cancer) and combination therapies [53, 54], is tailored specifically to fight the individual's cancer, knowing its genetic signature. Personalizing cancer therapy will also involve biomarker studies to understand the subpopulation of patients who may be the best responders to the drug. This, in turn, implies that the target population of each drug will become increasing more specific, and its testing will involve conducting multi-site, highly global trials to enable recruiting an adequate number of trial subjects. Innovative therapies of the future could include therapies that inhibit telomerase activity so that cancer cells cannot divide indefinitely [55, 56], or therapies that target cancer stem cells – cancer stem cells, which are tumorigenic, are found in certain tumors and hematologic cancers, and targeting them may prevent metastasis and lead to tumor inhibition [57, 58]. Novel delivery strategies for cancer therapeutics that have been developed or are being developed are resulting in or may result in more effective treatments [59]. All this could allow many cancers to morph from a deadly disease to a chronic manageable one sometime in the foreseeable future.

Figure 3 Clinical Trials Process (taken from http://www.mdanderson.es/sites/default/files/editor/elcancer/TRATAR/Trials%20Process%20ING.jpg)

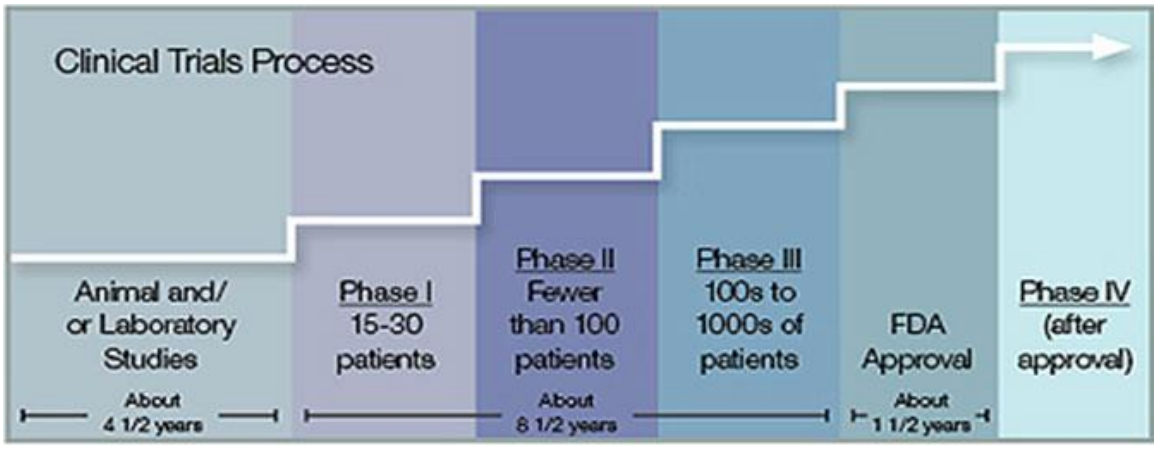
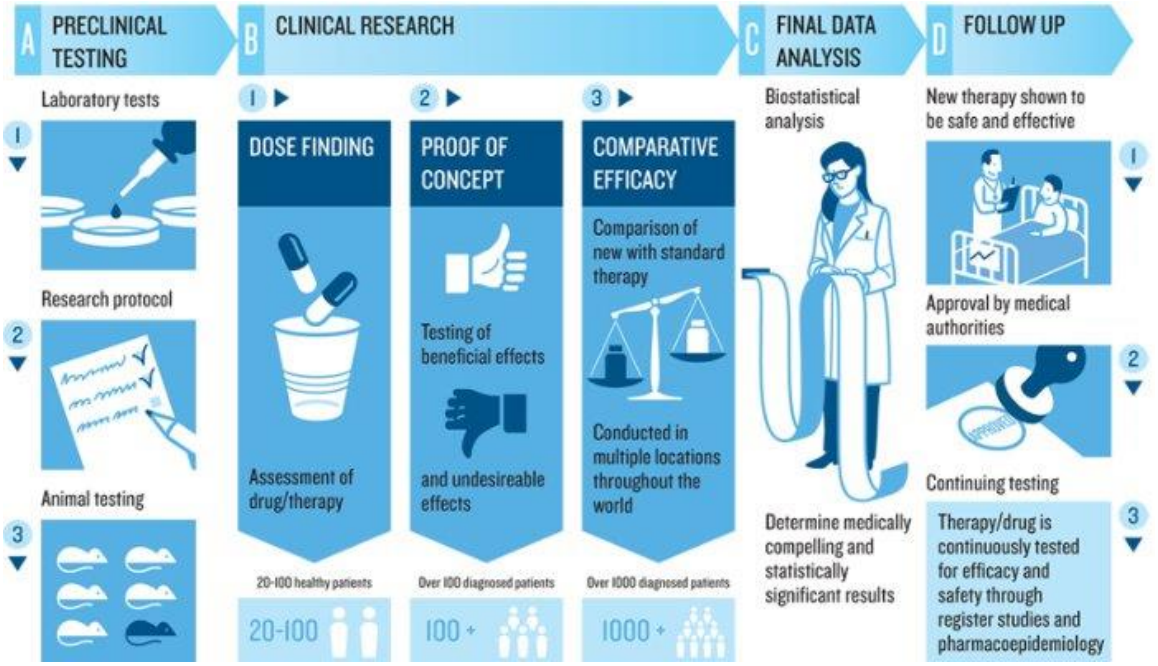


Figure 4 Clinical Trials Process (taken from https://www.linkedin.com/pulse/significance-process-evaluation-clinical-trials-joshua-ebenezer)



Aims and Scope of this Thesis

This statistical thesis mainly focuses on the design, statistical operating characteristics and interpretation of early phase oncology trials. While the later phase oncology trials, i.e. Phase III, typically have statistical endpoints such as overall survival and progression free survival, these endpoints are not the focus of early phase oncology trials. These early phase trials focus on finding the right safe dose to be used in the later stage trials. Due to the toxic nature of anti-cancer drugs and due to ethical concerns, cancer patients who have exhausted standard treatment options are generally recruited for this purpose, in contrast to many other therapeutic areas where healthy volunteers can be used in early phase trials. The high toxicity of these anti-cancer agents makes it imperative to minimize the possibility of giving patients unsafe or fatal doses. At the same time, it is also crucial to minimize the probability of giving patients sub-therapeutic or inefficacious doses. Thus, a careful study of drug toxicity or dose limiting toxicities (DLTs) and a determination of the maximum tolerated dose (MTD) is the focus of most Phase I oncology trials.

In this thesis, we first carefully study the statistical operating characteristics, including accuracy of MTD selection, percentage of patients assigned to the MTD, over-dosing, under-dosing, trial DLT rate, of eleven rule-based and model-based Phase 1 oncology designs that target or pre-specify a DLT rate of ~0.2 for three sets of true DLT probabilities. These DLT probabilities are generated at common dosages from specific linear, logistic, and log-logistic dose-toxicity curves. The simulations

performed in this thesis are intended to provide considerable information on design property trade-offs, and the means to explore additional settings. These simulations include studying the performance of the designs for different target DLT rates. Not every design allows specifying the target DLT rate – for those designs where the target DLT rate can be specified, the rate specified should depend on the severity of the disease. Designs where the target DLT rate cannot be changed should also be selected with the disease severity in mind. Previous works compare a very limited number of specific Phase 1 oncology designs. We fill this gap by comparing the performance of several designs that target the same DLT rate for the same true underlying DLT probabilities, and provide a practical aid in choosing a Phase I design for a particular setting or for developing a new Phase I oncology design.

Phase III oncology trials typically dose patients at the MTD that is determined from the corresponding Phase I trial. However, the sample size of the corresponding Phase I oncology trial is very limited and it is difficult to accurately predict the MTD with such a small number of patients. Using the wrong dose for safety and/or efficacy in a Phase III trial can have serious consequences for the development of the oncology drug in terms of cost, time and resources. Thus, it may be worthwhile to accurately determine the optimal drug dose for safety and efficacy in an early phase trial itself. Hence, we next study the effect of cohort size and sample size on the accuracy of dose selection in early phase oncology designs. We then propose a new design with a larger sample size that encompasses the objectives of both safety and efficacy and is easier

to implement than the existing designs. Designs such as the seamless Phase 1/2 SEARS design, a seamless 2-step Phase 1/2 design, designs to find the optimal biological dose and the Eff-Tox design among others have been proposed earlier to evaluate both drug toxicity and efficacy as an alternative to the standard approach of a Phase 1 followed by a Phase 2 trial. Instead, we propose the 20+20 accelerated titration design, a simple rule-based integrated Phase 1/2 trial design that selects an optimal dose for toxicity and efficacy via Bayesian decision rules at the end. We incorporate stopping rules within dose levels in this design to allow more flexible decision-making. In our simulations, we generate the true DLT rates at each dose based on a logistic dose-toxicity curve and the true response rates based on a manual dose-response curve since the true response rate of an anti-cancer drug may not always increase with an increase in dose. We finally compare the accuracy of optimal dose selection of this design with that of other Phase 1/2 strategies such as the Eff-Tox design, Optimal Biological Dose Isotonic design as well as a 3+3 Phase 1 design followed by a Phase 2 design.

We have also extended the TEQR and mTPI dose-finding oncology designs to choose an optimal dose for both safety and efficacy by considering correlated Bernoulli distributions for the true underlying toxicity and efficacy rates. In our simulations, we assume that the true DLT rates increase monotonically with an increase in dose but do not assume that the true response rates increase monotonically with an increase in dose; we allow multiple types of curves for dose response (monotonically

increasing, plateau, or umbrella-shaped). In this context, we apply isotonic regression to determine a dose that is optimal for safety and efficacy. Finally, we compare the accuracy of optimal dose selection of the extended TEQR and mTPI designs with that of the Eff-Tox design and of the Optimal Biological Dose Isotonic design.

In summary, through the work in this dissertation, we:

a) have obtained a comprehensive understanding of the statistical operating characteristics of several commonly used early phase oncology trial designs.
b) have understood the effect of sample and cohort size on the accuracy of dose selection in some early phase oncology designs.
c) have proposed some new early phase oncology designs with a larger sample size that consider both toxicity and efficacy in dose selection and that are improvements over the existing early phase designs.

It is thus hoped that the statistical characterization of early phase oncology designs has been furthered by this dissertation.

CHAPTER TWO[3]

Introduction

Phase I trials of a new anti-cancer drug are usually single arm, open label studies conducted on a small number (10s) of cancer patients, many of whom do not respond any longer to the standard treatment. Due to the toxic nature of many anti-cancer drugs as well as due to ethical reasons, cancer patients are enrolled in Phase I oncology trials, as opposed to the healthy volunteers used in Phase I trials in other therapeutic areas.

The main aim of a Phase I oncology trial is to investigate and understand the toxic properties (safety) of the new anti-cancer drug; the drug's efficacy is not traditionally the focus, although the drug's efficacy is often observed and monitored by the oncologist. With regard to safety, the trial helps investigators determine the right dose and dosing interval as well as the best route of administration of the new drug. In order to determine the right dose, an endpoint such as Phase 1 dose limiting toxicities (DLTs) in the first cycle is often considered.

For each dose finding Phase I trial, a set of pre-defined adverse events, typically only those possibly related to taking the study drug, constitutes the DLTs for that trial. Patients are traditionally monitored for DLTs during the first cycle of administration

[3] This chapter is published in Contemporary Clinical Trials Communications.

of the new anti-cancer drug; however, more recent trials may monitor DLTs for a longer period and may include toxicities in the DLT definition that are not included in the conventional definition of DLTs [60]. The starting dose in these dose finding trials is usually a very conservative dose based on animal studies of the drug, and the subsequent increasing doses to be administered are pre-specified. The number of patients with DLTs in each dose level is used to determine the Maximum Tolerated Dose (MTD). For a single anti-cancer drug being tested, the MTD is usually the highest dose level at which the observed DLT rate is equal to or below a specified percent. Phase II patients are generally dosed at the MTD determined in the corresponding Phase I trial. The above method for MTD selection is more applicable to cytotoxic agents where the toxicity and efficacy are assumed to increase monotonically with dose than to some modern molecularly targeted therapies where the MTD may not be reached even at higher doses due to their low toxicity; in such cases, another appropriate dosing endpoint may need to be considered such as the dose at which the key pharmacokinetic and pharmacodynamics parameters are optimal [60, 61].

Dose finding Phase I oncology designs can be broadly categorized [62, 63, 64, 65] as rule based (such as the 3+3 design) or model based (such as the CRM [66] and Eff-Tox designs [67]). The 3+3 design has been the workhorse dose finding design for Phase I oncology trials for a long time. It is still commonly used due to its simplicity and ease of implementation. However, depending on the target DLT rate of interest, it can be slow and inaccurate in estimating the MTD and can lead to a large portion of

patients receiving sub-therapeutic doses that do not produce any clinically meaningful response [68]. Hence, other designs, including model-based designs, have been explored in recent years [69, 70, 71].

The establishment of the MTD for various Phase 1 oncology designs is the main focus of this paper. In this work, we explore extensions of the 3+3 design as well as the model based mTPI [72], TEQR [73], BOIN [74], CRM [66, 62] and EWOC [75, 76] designs and compare their performance. There is no unique criterion to evaluate these designs since the performance of each design depends on the true DLT probability at each dose and the target DLT rate of the design. Hence, we systematically compare several statistical operating characteristics for the true DLT rates generated at the same doses by three different dose-toxicity curves. In addition, we explore the effect of starting the trial at different dose levels below the true MTD on the accuracy of MTD selection in these designs. The 3+3 design and its extensions we consider target a DLT rate of ∼0.2, and we specify a target DLT rate of 0.2 for the model based designs we consider. Although the results in this paper focus on a target DLT rate of 0.2, we explain in the discussion section the implications of targeting other DLT rates such as 0.1 and 0.33 with the A+B designs considered and discuss other A+B designs that target these rates. We also study the performance of the model based designs considered when the target DLT rate specified is 0.1 and 0.33. In contrast to previous works that compare a limited number of specific designs [77],

our comprehensive comparison across several designs should serve as a practical aid in applying these Phase I oncology designs or in developing new ones.

Methods

Rule Based Designs

We consider the 3+3 design, which targets a DLT rate of ∼0.2 [78], as well as its various extensions that target a DLT rate of ∼0.2. We also include the simple accelerated titration design and the 3+3+3 design in our study (Table 1) [79, 80, 81]. We then investigate several of their statistical operating characteristics, such as the accuracy of MTD selection among others. The formal definition of the MTD is that it is the dose for which Probability(DLT|dose=d)=target probability.

For the A+B designs [82] that allow only escalation, the algorithm that we follow is [80]:

1) If out of A patients assigned to dose level i, the number of DLTs observed is ≤x, then assign A patients to dose level i+1.
2) If the number of DLTs observed out of A patients at dose level i is >x and <y, then assign B more patients to dose level i. If out of A+B patients, the number of DLTs observed is ≤z, then add A patients to dose level i+1. Otherwise stop the trial.
3) If the number of DLTs observed out of A patients at dose level i is ≥y, then stop the trial.

We then estimate the MTD to be the dose level immediately below the last dose level examined. For the standard 3+3 design (Table 1), which is a special case of the general A+B design, this implies that the MTD is estimated to be the highest dose in which fewer than 33% of patients experience a DLT.

For the A+B designs that also allow de-escalation, the algorithm that we follow is:

1) Implement the rules given above for the corresponding escalation only design and let i be the dose level where the number of DLTs exceeds that allowed by the design. Then, ensure that A+B patients have been dosed at dose level i-1. If yes, dose level i-1 is estimated to be the MTD.
2) If not, add B more patients at dose level i-1.
    a) If out of the A+B patients at dose level i-1, the number of DLTs observed is ≤z, then dose level i-1 is estimated to be the MTD even if A+B patients have not been dosed at dose level i-2.
    b) If out of the A+B patients at dose level i-1, the number of DLTs observed is >z and A+B patients have been dosed at dose level i-2, then dose level i-2 is estimated to be the MTD. If A+B patients have not been dosed at dose level i-2, then add B more patients and continue the process.

For the 3+3 design with de-escalation, the MTD is estimated to be the highest dose in which fewer than 33% of patients experience a DLT, and in which at least six participants have been treated with the study drug.

For the rule-based designs where no de-escalation is allowed, Table 1 describes the dose finding rules; the specific x, y, and z for each A+B design can be determined based on the description of these designs in Table 1. To provide a preliminary idea of the properties of these designs, we depict in Figure 5 the probability of not escalating for a single step for various true DLT rates for the escalation only designs considered. For example, for the 3+3 design that allows only escalation, we can escalate at each step or dose level if 1) 0 out of 3 patients experience a DLT or if 2) 1 out of 6 patients experiences a DLT; the probability of escalating at each step or dose level is q^3+3pq^5 and not escalating at each step is $3p^2q+p^3+9p^2q^4+9p^3q^3+3p^4q^2$, where p is the probability of experiencing a DLT at the current dose level and q=1-p. Using these two probabilities and extending the framework to any number of steps, we can then calculate analytically the probability of selecting any dose level as the MTD for the 3+3 as well as other A+B designs that allow only escalation (see Lin, 2001 [83] and Appendix Table 1). This reference [83] also provides analytic formulae for the probability of MTD selection for the 3+3 and other A+B designs that allow de-escalation as well.

Model Based Designs or Designs that Allow Specification of the Target DLT rate

In terms of model-based designs, we consider the Modified Toxicity Probability Interval (mTPI), Toxicity Equivalence Range (TEQR), Bayesian Optimal Interval Design (BOIN), Continual Reassessment Method (CRM) and Escalation with Overdose Control (EWOC) designs and explore their statistical operating characteristics. For

these designs, we can choose the DLT rate that each design will target; we specify a target DLT rate of 0.2 for all of them, in order to compare their performance with the performance of the 3+3 design and its extensions that target a DLT rate of 0.2. Note that although the TEQR design is not a model based design, it allows the specification of the target DLT rate.

The mTPI design is described in detail in the reference by Ji and others [72]. The mTPI design is a Bayesian dose finding design that uses the posterior probability in guiding dose selection. The mTPI design uses a statistic for the decision rules called the unit probability mass (UPM), defined as the ratio of the probability mass of the interval and the length of the interval [72]. The toxicity probability scale is divided into three portions: $(0, p_T-\varepsilon_1)$ corresponding to under-dosing, $[p_T-\varepsilon_1, p_T+\varepsilon_2]$ corresponding to proper dosing and $(p_T+\varepsilon_2, 1)$ corresponding to over-dosing. Here p_T is the target probability of dose limiting toxicity and ε_1 and ε_2 are used to define the interval for the target DLT rate. The rules for escalating, staying at the same dose or de-escalating depend on which of these portions has the highest UPM for that dose level, based on a beta-binomial distribution with a beta(1,1) prior [72, 73]. For example, the next cohort of patients will be treated at the next higher dose level if the UPM is the largest for the under-dosing interval. The trial stops if dose level 1 is too toxic or if the maximum sample size is reached or exceeded.

The TEQR design is a frequentist version of the mTPI design and is described in detail

in the reference by Blanchard and Longmate [73]. This design is not based on the posterior probability but on the empirical DLT rate. The unit interval is divided into three portions: $(0, p_T-\varepsilon_1)$, $[p_T-\varepsilon_1, p_T+\varepsilon_2]$ and $(p_T+\varepsilon_2,1)$. The rules for escalating, staying at the same dose or de-escalating depend on which of these portions contains the empirical DLT rate for that dose level – if the empirical DLT rate lies between 0 and $p_T-\varepsilon_1$, we escalate; if it lies in the interval $[p_T-\varepsilon_1, p_T+\varepsilon_2]$, we stay at the same dose; if it lies above $p_T+\varepsilon_2$, we de-escalate. In both the mTPI and TEQR design, we stay at the current dose if the current dose is safe but the next higher dose is too toxic based on the data. A trial using the TEQR design stops if dose level 1 is too toxic or when a dose level achieves the selected MTD sample size. In a trial using the TEQR or the mTPI design, the MTD is determined to be the highest dose level with a DLT rate that is closest to (and below) the target DLT rate after applying isotonic regression at the end of the trial.

The concept of the BOIN design is similar to that of the TEQR design in terms of dividing the toxicity probability scale into three intervals and using these intervals along with the empirical DLT rate to guide dose finding [74]. In contrast to the TEQR and mTPI designs, where the interval for the target DLT rate is fixed and is independent of the dose level and the number of patients that have been treated at that dose level, the BOIN design is more general and permits this interval to vary with the dose level and the number of patients that have been treated at that dose level. In this design, the probability of patients being assigned to very toxic doses or to sub-

therapeutic doses is low. A trial using the BOIN design usually stops at the pre-planned sample size but the design allows the incorporation of early stopping rules.

The CRM design and its variations are well-known and are described in several references [84, 85, 86, 87]. This design uses the DLT information obtained from all the previous patients to determine the dose level to which the next patient (or cohort of patients [87]) is assigned. The first patient may be given a dose whose DLT rate is expected to be close to the target DLT rate based on information from previous studies, although caution usually dictates starting at a lower dose level. The dose given to each subsequent patient is decided by the DLT data of all the previous patients in conjunction with a dose-toxicity model for e.g. a one parameter logistic model with parameter "a". The estimates of "a" in the dose-toxicity model are updated using Bayesian methods after the DLT information from each patient is obtained. For example, after n patients are enrolled, $\widehat{a_n} = \int_0^\infty a\, f(a|\Omega_n)da$,

where $f(a|\Omega_n) = L_{\Omega_n}(a)g(a)/\int_0^\infty L_{\Omega_n}(a)g(a)da$;

$f(a|\Omega_n)$ is the posterior density of a, g(a) is the prior distribution for a, $L_{\Omega_n}(a)$ is the likelihood function, and Ω_n are the DLT data after n patients [88]. The dose-toxicity model is then used to recommend the dose level for the next patient, typically the dose with a DLT rate closest to but less than the updated DLT estimate from the model, subject to not skipping over untested doses. The stopping point for this process is usually the pre-determined sample size of the trial or an observation of no

change in dose assignment for a sequence of n patients.

The EWOC design is a Bayesian adaptive dose finding design, whose unique feature is over-dose control i.e. the posterior probability of treating patients at doses above the MTD, given the data, cannot be greater than a certain pre-specified probability α [75, 76]. In mathematical terms, we specify a prior distribution for (ρ_0, γ), where ρ_0 is probability of DLT at the minimum dose and γ is the MTD dose, and let $\Pi_n(\gamma)$ be the marginal posterior cdf of γ given D_n (DLT data after n patients). The first patient receives the dose x_1, and conditional on the event of no DLT at x_1, the $(n+1)^{th}$ patient receives the dose $x_{n+1}=\Pi^{-1}{}_n(\alpha)$, which implies that the posterior probability of exceeding the MTD is equal to α [76]. The design also minimizes the under-dosing of patients. This means that the MTD is generally reached rapidly, and after the initial cohorts of patients, the remaining cohorts of patients are treated at dose levels reasonably close to the MTD. In this design, it is also possible to add a stopping rule for excessive toxicity for e.g. the trial will be stopped early if three consecutive DLTs are observed or if the posterior probability at the minimum dose exceeds a certain pre-defined value.

Simulations of Rule Based Designs

For our simulations in SAS of the 3+3 design and its extensions, we use a Bernoulli random generator, along with the probability of a DLT at different doses generated

by a dose-toxicity curve, to randomly assign each patient a DLT or not depending on the probability of a DLT at the assigned dose. We then implement the assignment rules of each design and follow each simulated trial to its conclusion. For example, for the designs that allow only escalation, we escalate until the number of DLTs at the last dose level examined exceeds that allowed by the specific design, and the MTD is then estimated to be one dose level below the last dose level examined. We perform these simulations 10000 times for each combination of design and dose-toxicity curve. The increase in dose at a new dose level beyond dose level 1 for each dose-toxicity curve investigated is based on the modified Fibonacci series (2, 1.67, 1.5, 1.4, 1.33, 1.33, 1.33 etc.), as commonly used in many oncology trials [84].

A logistic dose-toxicity curve is often used to describe the underlying relation between dose and toxicity in cytotoxic agents [81]. Hence, we specify the true DLT probability at each dose based on a specific logistic curve. In addition to the logistic curve, we consider a specific log logistic and a linear dose-toxicity curve to study the performance sensitivity of these designs to the true DLT probabilities generated by these different dose-toxicity curves. Table 2 shows the true DLT rates at each dose level for each of the three dose-toxicity curves. For determining the two unknown coefficients of each dose-toxicity curve, we use the DLT rates at two different doses – namely we assume a true DLT rate of 0.01 at dose level 1 of 100 units and a DLT rate of 0.2 at the true MTD (dose level 3) of 334 units. We assume a DLT rate of 0.2 at the MTD because the 3+3 design targets a DLT rate between 0.2 and 0.25 [78]. Hence this

choice of 0.2 allows a fair comparison of the simulation results from the 3+3 design with those from other A+B designs whose approximate target DLT rate is 0.2 (various A+B designs target DLT rates other than 0.2; see Table 4.1 of Chapter 4 in the reference by Ting [89]). However, we also study the performance of these designs to different target DLT rates, such as 0.1 and 0.33.

We choose the following broad range of statistical operating characteristics to compare and evaluate the dose finding schemes considered for these three dose-toxicity curves: the accuracy of MTD selection, the average number of dose levels examined and its standard deviation, the maximum and median number of dose levels examined, the mean and median number of patients and the median number of DLTs per trial, the mean number of patients dosed at the MTD, the mean percentage of patients dosed at the MTD, above the MTD and below the MTD, the average number of patients and DLTs at each dose level, the average trial DLT rate and the average DLT rate at the MTD. Further, we investigate the effect of the location of the starting dose relative to the true MTD on the accuracy of MTD selection for the chosen logistic and log-logistic dose-toxicity curves for e.g. when we start our trial simulation at dose level -3, -2 or -1 instead of at dose level 1 (see Table 2; these low doses double each time). In addition, we use three linear dose-toxicity curves with different offsets to investigate the effect of the location of the starting dose relative to the true MTD on the accuracy of MTD selection for the 3+3 design. Our SAS programs, available on request, are presently able to provide results for six designs (3+3, 2+4, 4+4 a, 5+5 a,

3+3+3, and simple accelerated titration designs) and three dose-toxicity curves (linear, logistic, log-logistic). However, the programs are simple and flexible and can be extended to other A+B designs as well as any other dose-toxicity curve.

Simulations of Model Based Designs or Designs that Allow Specification of the Target DLT Rate

We use R code provided by Ji et al. [72] to implement the mTPI design. The program requires the following inputs: number of simulations, target probability of dose limiting toxicity p_T and ε_1 and ε_2 that help define the lower and upper bound of the interval for the target DLT rate respectively, sample size, cohort size, starting dose and the true DLT rate at each dose.

We use the R package TEQR to implement the TEQR design. The program requires the following inputs: number of simulations, target probability of dose limiting toxicity p_T and ε_1 and ε_2 that help define the lower and upper bound of the interval for the target DLT rate respectively, DLT probability deemed to be too toxic, desired sample size at the MTD, cohort size, maximum number of cohorts, starting dose and the true DLT rate at each dose.

We use the R package BOIN to implement the BOIN design. The program requires the following inputs: number of simulations, target probability of dose limiting toxicity

p_T, cohort size, number of cohorts, starting dose, cut off to eliminate an overly toxic dose for safety and the true DLT rate at each dose. Although the design allows the possibility of rules for stopping prior to reaching the planned sample size, we did not implement these early stopping rules, to permit fair comparisons between designs.

We use a CRM trial simulator to implement the various scenarios for the CRM design. The program requires the following inputs: maximum sample size, cohort size, number of doses, starting dose, target probability of dose limiting toxicity, stopping probability (the trial is stopped if the probability that the lowest dose is more toxic than the target is greater than this value) and the true DLT rates at the various doses. The probability of DLT at dose i is modeled as $p_i^{\exp(\alpha)}$, where p_i is a constant and α is distributed a priori as a normal random variable with mean 0 and variance 2. The initial default prior probabilities of DLT used in the software are given in Appendix Table 3. The trial stops when the planned sample size is reached or if the lowest dose is too toxic.

We use a web based program to implement the EWOC design. The program requires the following inputs: target probability of dose limiting toxicity, maximum acceptable probability of exceeding the target dose (α), variable α increment, cohort size, sample size, minimum dose, maximum dose, number of dose levels and the true probability of DLT at each dose. Although the EWOC design allows the possibility of rules for stopping prior to reaching the planned sample size, the current implementation of the

EWOC design does not include early stopping rules.

The parameters used for mTPI, TEQR, BOIN, CRM and EWOC designs are shown in Appendix Tables 2, 3, 4 and 5. Note that the sample size is an output of the rule-based A+B designs as well as the TEQR design. For the mTPI, BOIN, CRM and EWOC designs, we use the same sample size that the TEQR design yields for each of the three sets of true DLT rates.

Results

Comparison of Operating Characteristics for Designs that Target or Specify a DLT rate of 0.2

For all the simulation results in this section, dose level 1 is the lowest dose (see Table 2) and dose level 3 is the true MTD.

For the logistic dose-toxicity curve constructed, there is a very clear separation between the true DLT rate at the MTD and the rates at the dose levels below and above it: the DLT rate of 0.2 at the MTD versus 0.04 at the dose level below and 0.71 at the dose level above (Table 2). The DLT rate of 0.2 at dose level 3 aligns with the range of toxicity rates that the escalation-only A+B designs target (Table 1) and is the target DLT rate specified for the model-based designs. Hence all the designs pick dose level 3 as the MTD the largest percentage of times in our simulations, while incorrectly picking the other dose levels substantially less frequently (Table 3; also

see Appendix Table 1 for exact analytic results for MTD selection for the 3+3 design and its extensions). The 4+4a design with and without de-escalation, the mTPI design, the CRM design and the 3+3+3 design correctly pick dose level 3 as the MTD ~79%, ~80%, ~76%, ~76% and ~76% percent of the time respectively (Table 3 and Figure 6). The median number of patients enrolled in the trial ranges from 6 for the simple accelerated titration design to 25 for the 5+5 a design. As expected, with the 3+3 design, about half of the patients are given doses below the MTD. The BOIN design and the 5+5 a design with and without de-escalation also treat a large percentage of patients at doses below the MTD – about 50%, 48% and 49% respectively. On the other hand, the simple accelerated titration design over-doses a large percentage of patients (~43%). The model based designs generally treat a large percentage of patients at the MTD. The average trial DLT rate ranges from 0.17 for the TEQR design to 0.4 for the simple accelerated titration design; the median number of DLTs per trial ranges from 2 for the 2+4 design without de-escalation to 5 for the 4+4a design with de-escalation and the 5+5 a design, among the extensions of the 3+3 design considered.

For the log-logistic dose-toxicity curve constructed, there is a clear separation between the true DLT rate at the MTD and the rates at the dose levels below and above it: the DLT rate of 0.2 at the MTD versus 0.06 at the dose level below and 0.42 at the dose level above (Table 2). Although this separation is not as large as it is in the logistic dose-toxicity curve considered, all the designs still pick dose level 3 as the

MTD more frequently than they pick any other dose level. The CRM, mTPI, BOIN and 5+5 a with and without de-escalation designs correctly pick dose level 3 as the MTD ~74%, ~63%, ~59%, ~58% and ~58% percent of the time respectively (Table 4). The median number of patients enrolled in the trial ranges from 7 for the simple accelerated titration design to 30 for the 5+5 a design with de-escalation. For this dose-toxicity curve, about 49% of patients are given doses below the MTD in the 3+3 design. The BOIN, TEQR and 5+5 a design with and without de-escalation also treat a large percentage of patients at doses below the MTD – about 50%, 47%, 47% and 47% respectively. On the other hand, the simple accelerated titration design over-doses a large percentage of patients (~47%). The model based designs generally treat a large percentage of patients at the MTD. The average trial DLT rate ranges from 0.17 for the TEQR design to 0.34 for the simple accelerated titration design; the median number of DLTs per trial ranges from 2 for the simple accelerated titration design, reflecting the very small sample size for this design, to 5 for the 4+4 a design and the 5+5 a design with de-escalation, among the extensions of the 3+3 design considered.

For the linear dose-toxicity curve constructed, the DLT rate at dose level 3 is 0.2 and the DLT rate at dose level 4 is 0.34 (Table 2). Although this separation is even smaller than that in the logistic and log-logistic dose-toxicity curves considered, all the designs except the accelerated titration design (which picks dose level 3 as the MTD 27% of the time versus dose level 4 as the MTD 29% of the time) pick dose level 3 as the MTD more frequently than any other dose level. The CRM, mTPI, 5+5 a with and

without de-escalation and TEQR designs correctly pick dose level 3 as the MTD but only ~54%, ~45%, ~45%, ~45% and ~45% percent of the time respectively (Table 5). The median number of patients enrolled in the trial ranges from 8 for the simple accelerated titration design to 30 for the 5+5 a design with de-escalation. For this dose-toxicity curve, about half of the patients are given doses below the MTD in the 3+3 design. The BOIN, TEQR, CRM, mTPI designs and the 5+5 a design with and without de-escalation also treat a large percentage of patients at doses below the MTD – about 58%, 50%, 50%, 48%, 48% and 48% respectively. On the other hand, the simple accelerated titration over-doses a large percentage of patients (~49%). The model based designs generally treat a large percentage of patients at the MTD. The average trial DLT rate ranges from 0.16 for the TEQR design to 0.31 for the simple accelerated titration design; the median number of DLTs per trial ranges from 2 for the simple accelerated titration design to 5 for the 4+4 a and 5+5 a designs, among the extensions of the 3+3 design.

Results for the accuracy of MTD selection for the 3+3 design for all the three dose-toxicity curves considered are presented in Figure 7; results for some of the other designs are presented graphically in Appendix Figures 1–3.

Effect of Starting the Trial at Lower Dose Levels on the Accuracy of MTD Selection

In the previous section, our simulations are started at dose level 1 for all the rule-based designs, and dose level 3 is the true MTD for all the designs. This means that it takes only two escalations from the starting dose to reach the true MTD in the escalation only designs. However, the accuracy of MTD selection could depend on where the starting dose is located relative to the true MTD, for example if it is located six dose levels below the true MTD versus two, because some dose finding designs may be slow to escalate while others may be fast to do so. Thus, we investigate the effect of starting at lower dose levels on the accuracy of MTD selection in the 3+3 design and its extensions that allow only escalation, using the logistic dose-toxicity curve in Table 2. We find that the number of patients on the trial and the percentage of patients who are under-dosed, both of which are outputs of the program for the rule-based designs, increase when we start at the lower doses, but the accuracy of MTD selection is largely unaffected for all these designs (Table 6). We find similar results for the model based designs. We also find similar results for the log-logistic dose-toxicity curve in Table 2 to those described for the logistic dose-toxicity curve. The result that the location of the starting dose relative to the true MTD does not affect the accuracy of MTD selection may not be surprising since the true DLT rates at dose level -1, -2 and -3 are very small for the logistic and log-logistic dose-toxicity curves used.

In general, the accuracy of MTD selection will be affected when the true DLT rates at

these lower dose levels are much greater than 0.01 (say 0.1). We have demonstrated this for the 3+3 design using three linear dose-toxicity curves with different offsets (see Appendix Table 8 and Appendix Figure 4). In practice, the starting dose of the trial is usually an extremely conservative estimate based on animal studies, and the DLT rates at the first few dose levels are expected to be very low[4]. In this case, the accuracy of MTD selection should not be affected even when the true MTD is several doses above the starting dose in the rule-based escalation only designs considered, and we can enroll patients at the same low starting dose for these designs.

Discussion

In this work, we have systematically compared via simulations the statistical operating characteristics of various Phase I oncology designs, namely the 3+3 design and its extensions that target a DLT rate of ~0.2 as well as the mTPI, TEQR, BOIN, CRM and EWOC designs with a pre-specified target DLT rate of 0.2, for three sets of true DLT rates (generated for the same doses from a specific linear, logistic and log-logistic dose-toxicity curve). Although this is not an exhaustive comparison of all the current Phase 1 oncology designs, we have covered multiple commonly used ones. The 3+3 design is very simple and easy to implement and hence is still commonly

[4] While this is generally true, there are cases where the true DLT rate at low doses may not be close to zero, such as the following: 1) Phase 1 dose-finding trials sometimes consider all causality DLTs 2) The phrase "adverse events possibly related to study drug" in the definition of a DLT is considered to be "adverse events related to study drug", and it is often difficult to conclude whether an adverse event is due to the disease or the study drug. 3) The Phase 1 trial escalates a new drug added to an existing regimen that has toxicities.

used. However, our simulations show, not unexpectedly, that it under-doses a large percentage of patients, and is also not the design that picks the MTD most accurately for any of the dose-toxicity curves examined, with or without de-escalation.

All the designs examined select the MTD fairly accurately when there is a clear separation between the true DLT rate at the MTD and the rates at the dose level immediately below and above it, as is the case for the DLT rates generated using the chosen logistic dose-toxicity curve. However, when this separation is small, as is the case for the DLT rates generated using the chosen linear dose-toxicity curve, the accuracy of MTD selection is much lower. The separations in these true DLT rates depend, in turn, not only on the functional form of the dose-toxicity curve but also on the investigated dose levels and the parameter set-up. The considered A+B designs with de-escalation generally pick the MTD more accurately than the corresponding escalation-only design for the true DLT rates generated using the chosen log-logistic and linear toxicity curves, but not for the logistic one. Some of the other rule based designs examined pick the MTD more accurately than the 3+3 design, depending on the true DLT rate at each dose. For example, the 5+5 a design is as accurate as the model based designs in picking the MTD for the true DLT rates generated using the chosen log logistic and linear dose-toxicity curves but requires enrolling a larger number of patients compared to the other designs considered (~30 patients) and under-doses a large percentage of patients (~48%) for these dose-toxicity curves. Among the designs investigated, the simple accelerated titration design over-doses a

large percentage of patients. Over-dosing of patients in oncology trials is an important issue that needs to be considered carefully in terms of study design since the toxicities at the higher doses can be very harmful to patients. The EWOC design explicitly takes this into consideration; in this design, one can control the expected proportion of patients receiving doses above the MTD by pre-specifying the maximum acceptable probability of exceeding the target dose. Although some model-based designs can be more difficult to implement than rule based designs, the model based designs studied, mTPI, TEQR, BOIN, CRM and EWOC designs, perform well and assign the maximum percentage of patients to the MTD, and also have a reasonably high probability (given the small sample size) of picking the true MTD. The results for the Bayesian designs such as mTPI and EWOC may depend on the choice of the prior distribution, but our simulations have used the default prior for the software implementations and we have not performed any sensitivity analyses by changing the prior distributions in these designs.

In our simulations, we assumed a true DLT rate of 0.2 at the MTD (dose level 3) because it has been shown that the standard 3+3 design targets a toxicity rate between 0.2 and 0.25 [78]. However, when a DLT rate of 0.1 is specified as the target DLT rate, the various A+B designs considered would not, in general, select the MTD accurately because 0.1 is not within their target range, and when a DLT rate of 0.33 or 0.4 at the MTD is assumed, A+B designs that target a higher DLT rate would pick the MTD correctly more often than the 3+3 design. For example, for the linear dose-

toxicity curve in Table 2, dose level 2 is the true MTD if the target DLT rate is 0.1. In this case and for the extensions of the 3+3 design considered, percentages for correct MTD identification for dose level 2 are lower than those for dose level 3 and range from 14% (accelerated titration design) to 29% (5+5 a with target range 0.2–0.25); percentage for 3+3 is 27% (target range 0.17–0.26). If we consider a 5+5 design that targets a DLT range of 0.1–0.15 (see Table 4.1 of Chapter 4 of the reference by Ting [89]), it selects dose level 2 as the MTD ~43% of the time, which is much higher than the percentages with which the 3+3 and the other A+B designs with a target DLT rate of ~0.2 select dose level 2 as the MTD (results for this 5+5 design are not included in any table). Dose level 4 is the true MTD if the target DLT rate is 0.33. If we consider the 4+4 b design (target range 0.38–0.44) and 5+5 b design (target range 0.3–0.35) (see Table 4.1 of Chapter 4 of the reference by Ting [89]), they both select dose level 4 as the MTD ~40% of the time (results not shown here). This is much higher than the percentages with which the 3+3 and the other A+B designs with a target DLT rate of ~0.2 select dose level 4 as the MTD for the chosen linear dose-toxicity curve (percentages for correct MTD identification range from 20% to 31%). Results for the accuracy of MTD selection for the model based designs for the linear dose-toxicity curve given in Table 2 and for the target DLT rates of 0.1 and 0.33 are provided in Appendix Tables 6 and 7 respectively. The accuracy of MTD selection decreases as the target DLT rate increases from 0.1 to 0.33 for the mTPI, TEQR, BOIN and CRM designs, but not for the EWOC design, for the chosen linear dose-toxicity curve. Our simulations for the A+B and model based designs show that for designs where the

approximate DLT rate targeted by the design is known, it is critical to pick a design that is aligned with the true DLT rate of interest.

We also showed that as long as the true DLT rates at the first few dose levels are very low, the accuracy of MTD selection is largely unaffected by the number of escalations it takes to reach the true MTD, for the rule-based escalation only designs considered that target a DLT rate of ~0.2.

For the standard 3+3 design, our simulations, where the starting dose is two levels below the true MTD, show that the maximum number of dose levels examined varies between 5 for the logistic dose-toxicity curve and 7 for the linear and log-logistic dose-toxicity curves considered, while the median number of dose levels examined is 4 for all the three dose-toxicity curves. In comparison, a literature review of 41 trials that were performed using the standard 3+3 design found that the median number of dose levels examined was 6 (range 2–12 dose levels), about 45% of the patients were under-dosed and about 20% of the patients were over-dosed [90]. These empirical results are consistent with our simulation findings that the 3+3 design under-doses about 50% of the patients and over-doses about 22% of the patients on the trial, for all the three dose-toxicity curves. The average number of patients enrolled in trials that are based on the 3+3 design is, however, much higher in the literature review with a mean of 44 patients than in our simulations, where we found a mean of ~14 patients for all the three dose-toxicity curves. However, this literature review is based

on trials of targeted anti-cancer agents that reached the MTD and we do not know the exact percentage of trials that included expansion cohorts, and if the initial cohorts started at very low doses; hence, the above comparisons are not exact. Nevertheless, it is clear from clinical trial data as well as our simulations that Phase I trials are very small and thus may not provide good estimates of the MTD. If we consider designs with a higher average sample size, say 50–60 patients, they will have a much higher accuracy of MTD selection. In the future, it may be worthwhile investing in the enrollment of a larger number of patients even in a Phase I trial to obtain more accurate estimates of the right dose to be used for later Phase trials, although there is always a trade-off between costs (lower number of patients) and more accurate estimates (higher number of patients).

Conclusions

In conclusion, our comprehensive study compares and contrasts the 3+3 design with multiple other Phase I oncology designs with an approximate target DLT rate of 0.2 for various scenarios of true underlying DLT rates, in order to understand which designs pick the true MTD most accurately, which under-dose and over-dose the maximum percentage of patients, which assign the maximum number and percentage of patients to the MTD cohort, which explore the maximum number of dose levels and enroll the most number of patients in each case. Our SAS programs are flexible and can be extended to include other A+B designs, other dose-toxicity curves as well as other evaluation criteria. The summaries in this paper provide considerable

information on design property trade-offs, and the means to explore additional settings. These may be useful aids in choosing a Phase I design for a particular setting.

Table 1 Designs Investigated that are Extensions of the 3+3 Design that Allow Only Escalation

Design	Assignment Rule	Ways to Escalate	Approximate Range for Toxicity Rate Targeted by the Design (Table 4.1, Chapter 4, Ting, 2006 [89]; Storer, 2001 [78])
3+3	If 0 out of 3 enrolled patients have a DLT, then escalate to the next dose level and enroll 3 more; if 1 out of 3 patients has a DLT, then add 3 more patients at the same dose level; if 2 or more patients out of 3 or 6 patients experience a DLT, then stop the trial. The MTD is one dose level below.	0/3 =0% or 1/6=16.7% i.e. can escalate if we observe 0 DLTs out of 3 patients, or 1 DLT out of 6 patients	$0.17<\Gamma<0.26$ or $0.2<\Gamma<0.25$
2+4	If 0 out of 2 enrolled patients have a DLT, then escalate to the next dose level and enroll 2 more; if 1 out of 2 patients has a DLT, then add 4 more patients at the same dose level; if 2 or more patients out of 2 or 6 patients experience a DLT, then stop the trial. The MTD is one dose level below.	0/2=0% or 1/6=16.7% i.e. can escalate if we observe 0 DLTs out of 2 patients, or 1 DLT out of 6 patients	$0.17<\Gamma<0.26$
4+4 a	If 0 out of 4 enrolled patients have a DLT, then escalate to the next dose level and enroll 4 more; if 1 or 2 out of 4 patients have a DLT, then add 4 more patients at the same dose level; if 3 or more patients out of 4 or 8 experience a DLT, then stop the trial. The MTD is one dose level below.	0/4=0% or 1/8=12.5% or 2/8=25% i.e. can escalate if we observe 0 DLTs out of 4 patients, or 1 DLT out of 8	$0.25<\Gamma<0.31$

		patients, or 2 DLTs out of 8 patients	
5+5 a	If 0 out of 5 enrolled patients have a DLT, then escalate to the next dose level and enroll 5 more; if 1 or 2 out of 5 patients have a DLT, then add 5 more patients at the same dose level; if 3 or more patients out of 5 or 10 experience a DLT, then stop the trial. The MTD is one dose level below.	0/5=0% or 1/10=10% or 2/10=20% i.e. can escalate if we observe 0 DLTs out of 5 patients, or 1 DLT out of 10 patients, or 2 DLTs out of 10 patients	$0.2<\Gamma<0.25$
3+3+3	If 0 out of 3 enrolled patients have a DLT, then escalate to the next dose level and enroll 3 more; if 1 out of 3 patients has a DLT, then add 3 more patients at the same dose level; if 2 out of 6 patients have a DLT then add 3 more patients at the same dose level; if 2 or more patients out of 3 patients experience a DLT or 3 or more out of 6 or 9 patients experience a DLT, then stop the trial. The MTD is one dose level below.	0/3=0% or 1/6=16.7% or 2/9=22.2% i.e. can escalate if we observe 0 DLTs out of 3 patients, or 1 DLT out of 6 patients, or 2 DLTs out of 9 patients	
Simple Accelerated Titration Design	Successively assign a single patient at each dose level until the patient has a DLT. Then switch to the 3+3 design (i.e. add 2 more patients to the dose level at which a DLT is first seen and then follow the rules of the 3+3 design).		

The table above provides the rules for the escalation only designs but we also allow de-escalation in the 3+3, 2+4, 4+4 a, and 5+5 a designs and follow the algorithm described in the methods section. The designs that also allow de-escalation will target a slightly lower DLT rate than their counterparts that allow only escalation. One method to estimate the approximate target DLT rate of each design that also allows de-escalation is to run simulations for each design using

several different dose-toxicity curves and then perform the following calculation: one needs to compute the sum of the product of the true DLT rate at each dose and the probability that that dose is selected as the MTD from simulations for each scenario and then find the average of this value across the various scenarios (dose-toxicity curves). Based on our results for the logistic, log-logistic and linear dose-toxicity curves in Table 3-5, we find that the approximate target DLT rate of the 3+3 design with de-escalation is 0.17, of the 2+4 design with de-escalation is 0.18, of the 4+4 a design with de-escalation is 0.21 (which is why we also included the 4+4 a design, even though its target DLT rate for the escalation only case is a little higher than 0.2), and of the 5+5 a design with de-escalation is 0.17. The 3+3+3 design targets an approximate DLT rate of 0.21.

Table 2 DLT Rates at Different Doses for the Three Dose-toxicity Curves (also see Appendix Figure 5)

		Linear Dose-toxicity DLT rate= min (-0.071197+0.000811966*dose,1)	**Logistic Dose-toxicity Log(DLT rate/(1-DLT rate)) =-5.96641+ 0.013713*dose**	**Log-Logistic Dose-toxicity Log(DLT rate/(1-DLT rate)) =-16.8485+2.66078*log(dose)**
Dose Level	**Dose**	**DLT Rate**	**DLT Rate**	**DLT Rate**
-3	12.5 units		0.00303	0.00004
-2	25		0.0036	0.0003
-1	50		0.00506	0.0016
1	100	0.01	0.01	0.01
2	200	0.09	0.04	0.06
3	334	0.2	0.2	0.2
4	501	0.34	0.71	0.42
5	701.4	0.50	0.97	0.64
6	932.86	0.69	1	0.79
7	1240.71	0.94	1	0.89
8	1650.14	1	1	0.95
9	2194.69	1	1	0.97
10	2918.93	1	1	0.99

Table 3 Simulation Results: Logistic Dose-toxicity: Log(DLT rate/(1-DLT rate)) =-5.96641+ 0.013713*dose

Design	% of times that dose level 3 is selected as the MTD	% of times that doses below the MTD (dose levels 1 and 2) are selected as the MTD	% of times that doses above the MTD (dose levels 4 and above) are selected as the MTD	Average Number of Dose Levels Examined	Std of Dose Levels Examined	Max Dose Levels Examined	Median Dose Levels Examined	Average Number of Patients per Trial	Median Number of Patients per Trial	Median Number of DLTs per Trial	Average Sample Size at MTD	Average % of pts dosed at MTD	Average % of pts under-dosed	Average % of pts over-dosed
3+3*	68.05 (64.32)	29.78 (34.75)	2 (0.76)	3.7	0.54	5	4	13.06 (15.53)	12 (15)	3 (3)	4.1 (5.6)	31.43 (35.87)	**50.30 (48.47)**	18.26 (15.66)
2+4*	69.62 (64.67)	23.54 (34.47)	6.77 (0.77)	3.8	0.56	5	4	10.48 (14.59)	10 (14)	2 (3)	3.22 (5.75)	30.86 (40.17)	43.23 (39.69)	25.90 (20.14)
4+4 a*	**79.65 (78.79)**	19.39 (20.45)	0.96 (0.75)	3.8	0.42	5	4	19.23 (21.63)	20 (20)	4 (5)	6.24 (7.86)	32.67 (36.68)	46.08 (44.13)	21.26 (19.19)
5+5 a*	69.19 (67.5)	30.68 (32.43)	0.13 (0.05)	3.7	0.47	5	4	23.14 **(26.12)**	**25 (25)**	5 (5)	8.05 (9.67)	34.96 (37.13)	**48.83 (48.26)**	16.21 (14.61)
3+3+3	**75.9**	21.77	2.3	3.8	0.47	5	4	13.96	15	3	4.59	32.25	47.72	20.03
Simple accelerated titration	62.98	14.51	22.43	4.1	0.64	6	4	**7.14**	**6**	3	1.88	24.90	32.06	**43.04**
mTPI	**76.1**	23	0.85	5		5	5	21 (max)	21 (max)		10.1	**47.88**	41.9	10.22
TEQR	70	27	1	5		5	5	21.78	21		9.5	**43.66**	46.74	9.6
BOIN	72.3	25.4	2.3	5		5	5	21 (max)	21 (max)	3.4 (mean)	8.6	41.15	**49.76**	9.09
CRM	**76**	21	3	5		5	5	21 (max)	21 (max)	3.6 (mean)	9.8	**46.88**	43.97	9.15
EWOC	70.45	9.7	19.85	5		5	5	21 (max)	21 (max)		10.1	**48.04**	40.06	11.9

The bold highlighting shows the designs predicted by simulations to pick the MTD most accurately, to enroll the largest and smallest number of patients, to dose the maximum percentage of patients at the MTD, to under-dose the maximum percentage of patients, and to over-dose the maximum percentage of patients. Note also that the sum of columns 2 to 4 may add up to <100% because the remaining small percentage of times, no dose level is selected as the MTD.

* The numbers shown in brackets are for a corresponding design that allows dose de-escalation.

Table 4 Simulation Results: Log-Logistic Dose-toxicity: Log(DLT rate/(1-DLT rate)) =-16.8485+2.66078*log(dose)

Design	% of times that dose level 3 is selected as the MTD	% of times that doses below the MTD (dose levels 1 and 2) are selected as the MTD	% of times that doses above the MTD (dose levels 4 and above) are selected as the MTD	Average Number of Dose Levels Examined	Std of Dose Levels Examined	Max Dose Levels Examined	Median Dose Levels Examined	Average Number of Patients per Trial	Median Number of Patients per Trial	Median Number of DLTs per Trial	Average Sample Size at MTD	Average % of pts dosed at MTD	Average % of pts under-dosed	Average % of pts over-dosed
3+3*	49.45 (50.55)	31.66 (35.95)	18.72 (13.38)	3.8	0.8	7	4	14.2 (16.73)	15 (15)	3 (3)	4.00 (5.18)	28.72 (31.16)	**48.61 (47.44)**	22.67 (21.4)
2+4*	45.8 (50.89)	24.48 (33.94)	29.6 (15.05)	4.1	0.87	8	4	11.89 (16.29)	12 (16)	3 (3)	3.16 (5.19)	27.49 (32.71)	40.05 (37.8)	32.46 (29.49)
4+4 a*	56.73 (57.76)	20.26 (20.69)	23.01 (21.54)	4	0.7	6	4	21.96 (24.23)	20 (24)	5 (5)	6.18 (7.4)	29.09 (31.3)	42.78 (41.49)	28.13 (27.21)
5+5 a*	**58.07 (58.09)**	31.38 (33.18)	10.54 (8.71)	3.8	0.65	6	4	25.54 **(28.43)**	25 **(30)**	4 (5)	7.96 (9.37)	31.95 (33.38)	**46.85 (46.63)**	21.21 (19.99)
3+3+3	53.96	22.43	23.56	4	0.74	7	4	15.89	15	3	4.55	28.9	44.54	26.56
Simple accelerated titration	36.32	15.67	47.95	4.5	1.05	9	4	**8.11**	**7**	2	1.87	22.93	29.81	**47.25**
mTPI	**63.15**	22.45	14.35	7		7	7	24 (max)	24 (max)		10.0	**41.67**	40.49	17.85
TEQR	57	32	8	7		7	7	22.71	24		8.6	**37.78**	**46.98**	15.24
BOIN	**59.2**	28	12.7	7		7	7	24 (max)	24 (max)	3.7 (mean)	8.9	**37.08**	**50**	12.92
CRM	**74**	18	8	7		7	7	24 (max)	24 (max)	4.0 (mean)	10.1	**41.92**	43.42	14.67
EWOC	57.1	9.7	33.2	7		7	7	24 (max)	24 (max)		11.4	**47.32**	22.92	29.76

The bold highlighting shows the designs predicted by simulations to pick the MTD most accurately, to enroll the largest and smallest number of patients, to dose the maximum percentage of patients at the MTD, to under-dose the maximum percentage of patients, and to over-dose the maximum percentage of patients. Note also that the sum of columns 2 to 4 may add up to <100% because the remaining small percentage of times, no dose level is selected as the MTD.
* The numbers shown in brackets are for a corresponding design that allows dose de-escalation.

Table 5 Simulation Results: Linear Dose-toxicity: DLT rate=min(-0.071197+0.000811966*dose, 1)

Design	% of times that dose level 3 is selected as the MTD	% of times that doses below the MTD (dose levels 1 and 2) are selected as the MTD	% of times that doses above the MTD (dose levels 4 and above) are selected as the MTD	Average Number of Dose Levels Examined	Std of Dose Levels Examined	Max Dose Levels Examined	Median Dose Levels Examined	Average Number of Patients per Trial	Median Number of Patients per Trial	Median Number of DLTs per Trial	Average Sample Size at MTD	Average % of pts dosed at MTD	Average % of pts under-dosed	Average % of pts over-dosed
3+3*	37.49 (39.86)	34.6 (37.62)	27.72 (22.39)	3.9	1.01	7	4	14.75 (17.22)	15 (18)	3 (3)	3.85 (4.76)	26.44 (27.73)	**49.64 (48.67)**	23.92 (23.60)
2+4*	34.59 (39.72)	26.88 (33.93)	38.42 (26.27)	4.2	1.1	7	4	12.52 (16.9)	12 (16)	3 (3)	3.08 (4.63)	25.52 (28.2)	40.75 (38.7)	33.73 (33.1)
4+4 a*	40.56 (41.94)	21.47 (21.68)	37.97 (36.36)	4.2	0.92	7	4	23.64 (25.78)	24 (24)	5 (5)	6.07 (6.97)	26.73 (27.96)	42.52 (41.28)	30.75 (30.76)
5+5 a*	**44.59 (45.44)**	33.92 (35.13)	21.48 (19.41)	3.8	0.85	6	4	26.85 **(29.63)**	25 **(30)**	5 (5)	7.66 (8.74)	29.24 (29.88)	**47.87 (47.9)**	22.89 (22.23)
3+3+3	39.56	24.73	35.63	4.1	0.97	7	4	16.99	18	3	4.43	26.57	44.55	28.89
Simple accelerated titration	26.69	16.99	56.26	4.7	1.26	8	5	**8.67**	**8**	2	1.85	21.5	29.94	**48.57**
mTPI	**45.3**	28.6	26.05	7		7	7	21 (max)	21 (max)		6.9	**32.71**	**47.99**	19.29
TEQR	**45**	37	15	7		7	7	22.88	21		7.4	**32.12**	**49.78**	18.09
BOIN	40.4	38.1	21.6	7		7	7	21 (max)	21 (max)	3.0 (mean)	6.1	29.05	**57.62**	13.33
CRM	**54**	24	22	7		7	7	21 (max)	21 (max)	3.3 (mean)	7.2	**34.43**	**49.57**	16.00
EWOC	40.35	8.90	50.75	7		7	7	21 (max)	21 (max)		8.5	**40.39**	23.81	35.81

The bold highlighting shows the designs predicted by simulations to pick the MTD most accurately, to enroll the largest and smallest number of patients, to dose the maximum percentage of patients at the MTD, to under-dose the maximum percentage of patients, and to over-dose the maximum percentage of patients. Note also that the sum of columns 2 to 4 may add up to <100% because the remaining small percentage of times, no dose level is selected as the MTD.

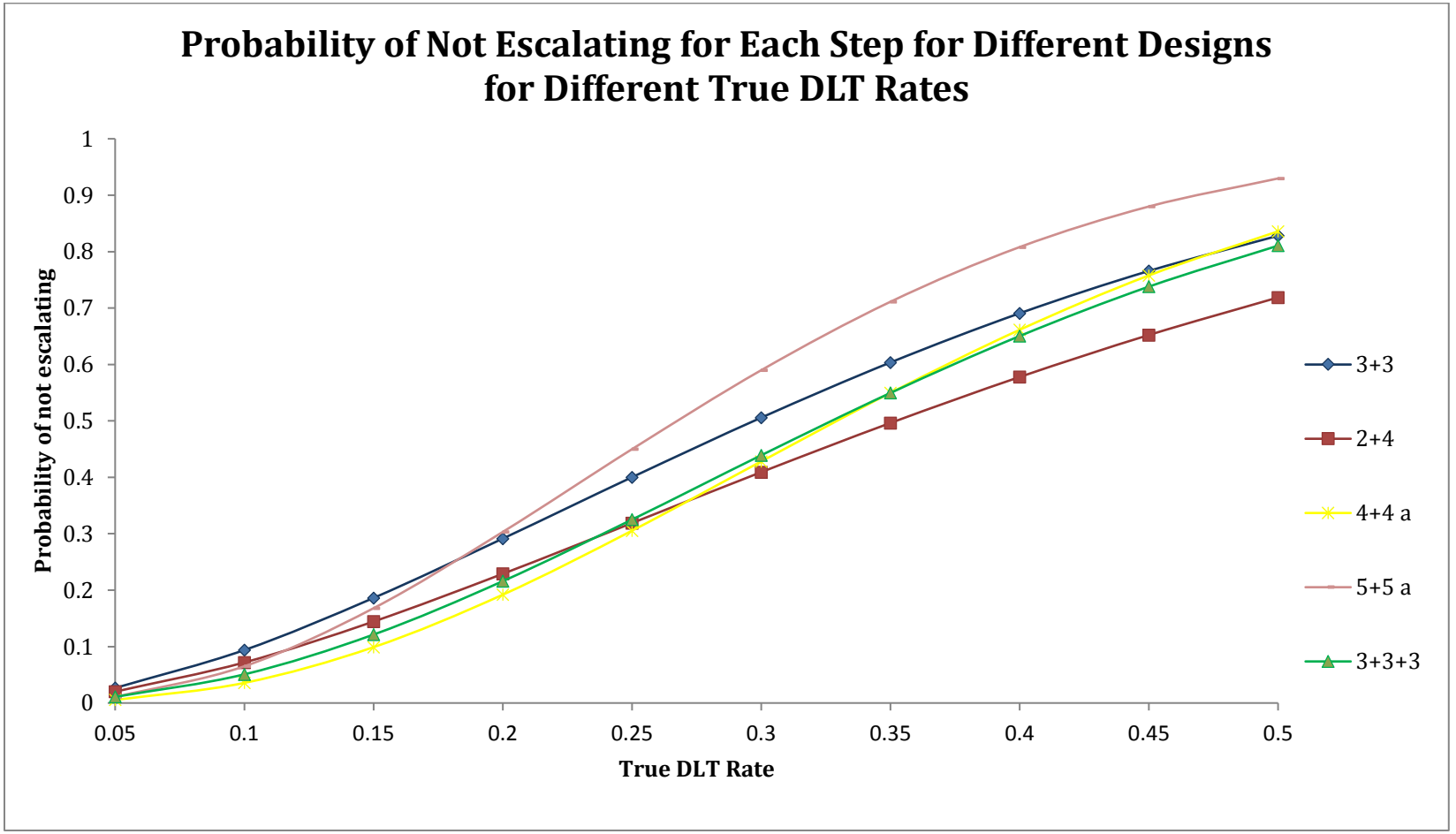
* The numbers shown in brackets are for a corresponding design that allows dose de-escalation.

Table 6 Simulation Results: Logistic Dose-toxicity: Log(DLT rate/(1-DLT rate)) =-5.96641+ 0.013713*dose: Effect of Starting at Lower Doses on the Accuracy of MTD Selection

Design	Median Sample Size when Starting Dose is Dose Level -3	Accuracy of MTD Selection when Starting Dose is Dose Level -3 (% of Times Dose Level 3 is Selected as MTD)	% of Patients Underdosed when Starting Dose is Dose Level -3	Median Sample Size when Starting Dose is Dose Level -2	Accuracy of MTD Selection when Starting Dose is Dose Level -2 (% of Times Dose Level 3 is Selected as MTD)	% of Patients Underdosed when Starting Dose is Dose Level -2	Median Sample Size when Starting Dose is Dose Level -1	Accuracy of MTD Selection when Starting Dose is Dose Level -1 (% of Times Dose Level 3 is Selected as MTD)	% of Patients Underdosed when Starting Dose is Dose Level -1	Median Sample Size when Starting Dose is Dose Level 1	Accuracy of MTD Selection when Starting Dose is Dose Level 1 (% of Times Dose Level 3 is Selected as MTD)	% of Patients Underdosed when Starting Dose is Dose Level 1
3+3	21	67.79%	70.84%	18	67.11%	66.16%	15	67.82%	59.74%	12	68.05%	50.30%
2+4	16	70.45%	63.96%	14	70.51%	59.06%	12	69.76%	52.45%	10	69.62%	43.23%
4+4 a	32	79.59%	67.02%	28	79.54%	62.21%	24	79.53%	55.66%	20	79.65%	46.08%
5+5 a	40	69.48%	69.18%	35	69.05%	64.48%	30	69.80%	58.22%	25	69.19%	48.83%
3+3+3	24	75.92%	68.45%	21	75.85%	63.66%	18	76.09%	57.14%	15	75.9%	47.72%
Accelerated Titration	9	63.35%	52.96%	8	63.79%	47.60%	7	63.00%	41.04%	6	62.98%	32.06%
mTPI	30 (max)	77.3%	59.71%	27 (max)	77.5%	54.49%	24 (max)	77.8%	48.72%	21 (max)	76.1%	41.9%
TEQR	30	70%	62.91%	27	69%	59.17%	24	71%	54.22%	21	70%	46.74%
BOIN	30 (max)	72.2	65.89%	27 (max)	71%	62.08%	24 (max)	72%	57.5%	21 (max)	72.3%	49.76%
CRM	30 (max)	75%	60.88%	27 (max)	76%	56.2%	24 (max)	76%	50.73%	21 (max)	76%	43.97%
EWOC	30 (max)	70.2%	52.76%	27 (max)	70.85%	45.76%	24 (max)	65.3%	47.4%	21 (max)	70.45%	40.06%

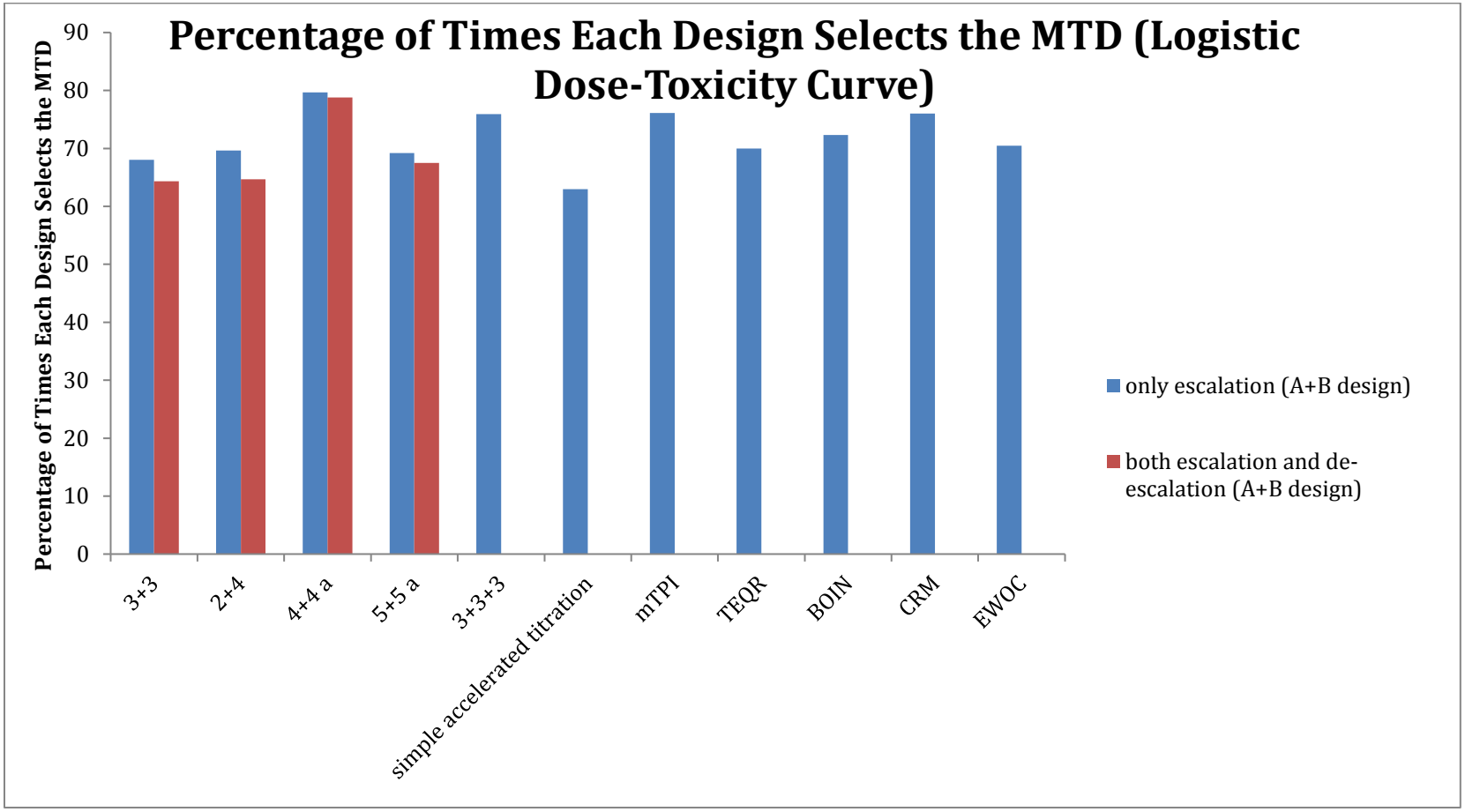
The sample size is an output for the A+B escalation only designs. For the model based designs, the sample size is an output for the TEQR design and we use the same sample size obtained from the TEQR design for the other model based designs. For the CRM design, a prior DLT rate of 0.01, 0.05 and 0.1 are used at dose levels -1, -2 and -3.

Figure 5 The probability of not escalating at each step for different true DLT rates for the escalation only designs considered that are extensions of the 3+3 design and that target a DLT rate of ~0.2.



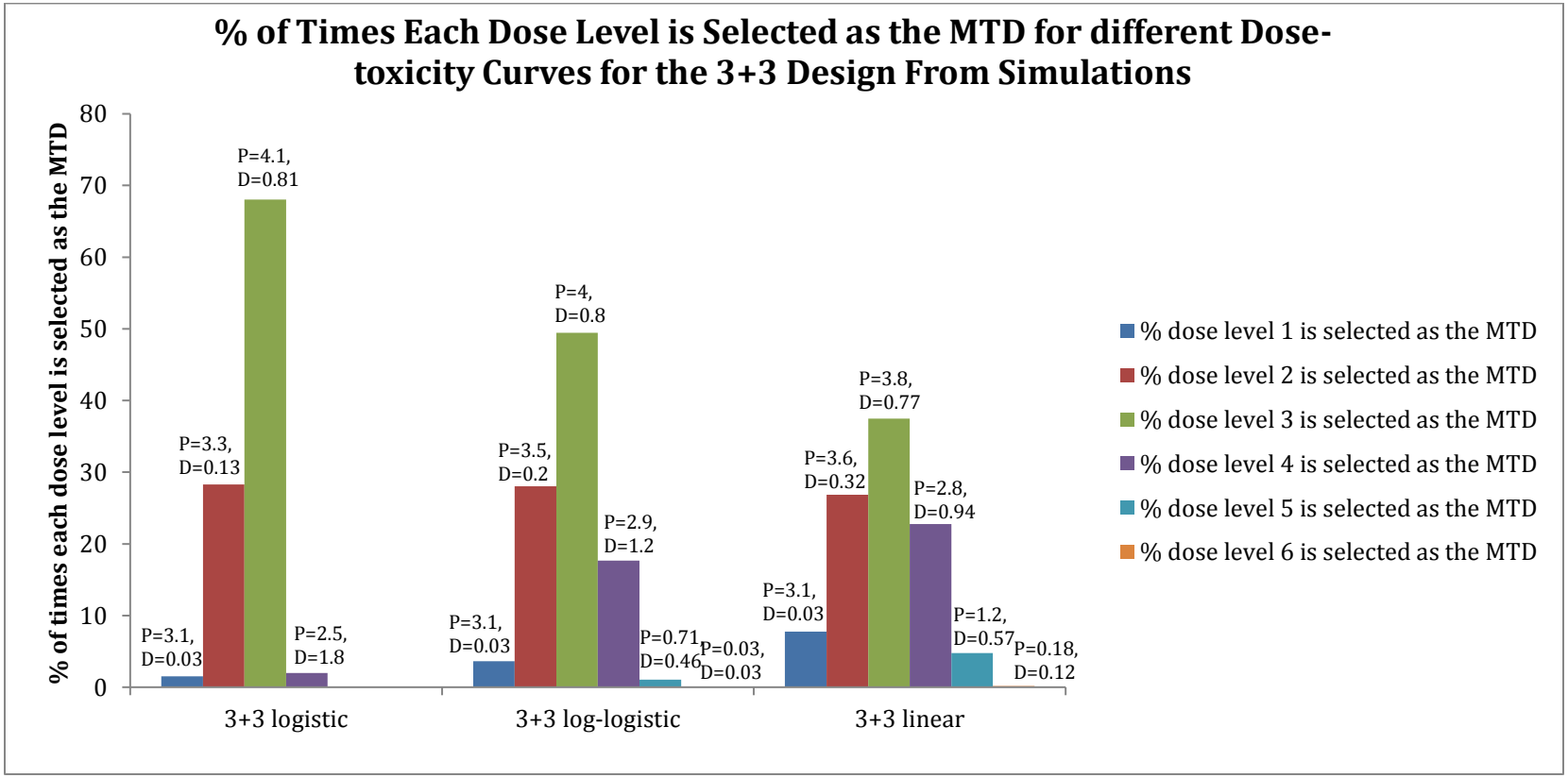
Legend: Figure 5 depicts the probability of not escalating at each step for different true DLT rates for the escalation only designs considered that are extensions of the 3+3 design and that target a DLT rate of ~0.2. These probabilities are derived analytically based on the decision rules of each design as given in Table 1.

Figure 6 The percentage of times each design considered selects the MTD (Dose Level 3) for the true DLT rates generated from the logistic dose-toxicity curve given in Table 2.



Legend: Figure 6 depicts the percentage of times each design considered selects the MTD (Dose Level 3) for the true DLT rates generated from the logistic dose-toxicity curve given in Table 2. These percentages are from simulations and the results are shown in Tables 3-5.

Figure 7 The percentage of times that the 3+3 design selects each dose level as the MTD for the true DLT rates given in Table 2.



3+3 logistic implies the 3+3 design with the DLT rates generated from the logistic dose-toxicity curve in Table 2 and similarly for the others.
P=Mean Sample Size and D=Mean Number of DLTs at each dose level (from 10000 simulations).

Legend: Figure 7 depicts the percentage of times that the 3+3 design selects each dose level as the MTD for the true DLT rates given in Table 2, generated from the three dose-toxicity curves. These percentages are from simulations and the results are shown in Tables 3–5.

Appendix Chapter 2

Appendix Table 1: Analytic Results for MTD Selection:
Logistic Dose-toxicity Curve:

Dose Level	Probability of DLT	Probability of Being the Highest Dose Level Examined				
		3+3	2+4	4+4 a	5+5 a	3+3+3
1	0.01	0.00	0.00	0.00	0.00	0.00
2	0.04	0.02	0.01	0.00	0.01	0.01
3	0.2	0.29	0.23	0.19	0.30	0.21
4	**0.71**	**0.68**	**0.70**	**0.79**	**0.69**	**0.76**
5	0.97	0.02	0.07	0.01	0.00	0.02

Log-Logistic Dose-toxicity Curve:

Dose Level	Probability of DLT	Probability of Being the Highest Dose Level Examined				
		3+3	2+4	4+4 a	5+5 a	3+3+3
1	0.01	0.00	0.00	0.00	0.00	0.00
2	0.06	0.04	0.03	0.01	0.02	0.02
3	0.2	0.28	0.22	0.19	0.30	0.21
4	**0.42**	**0.50**	**0.46**	**0.57**	**0.58**	**0.53**
5	0.64	0.17	0.25	0.22	0.10	0.22
6	0.79	0.01	0.04	0.01	0.00	0.01

Linear Dose-toxicity Curve:

Dose Level	Probability of DLT	Probability of Being the Highest Dose Level Examined				
		3+3	2+4	4+4 a	5+5 a	3+3+3
1	0.01	0.00	0.00	0.00	0.00	0.00
2	0.09	0.08	0.06	0.03	0.05	0.04
3	0.2	0.27	0.21	0.19	0.29	0.21
4	**0.34**	**0.38**	**0.34**	**0.40**	**0.45**	**0.39**
5	0.5	0.23	0.27	0.32	0.20	0.29
6	0.69	0.05	0.10	0.06	0.02	0.07
7	0.94	0.00	0.01	0.00	0.00	0.00

These are exact analytic results for MTD selection for extensions of the 3+3 design that allow only escalation, and the results are very close to those provided in Tables 3–5 for MTD selection, which are based on simulations. The rows highlighted in bold show the probability of dose level 3 being chosen as the MTD for the various designs and dose-toxicity curves.

Appendix Table 2: Parameters for the mTPI and TEQR Designs:

Parameter	mTPI Design	TEQR Design
Number of simulations	2000	2000
Target toxicity probability p_T	0.2	0.2
ε_1	0.05	0.05
ε_2	0.05	0.05
Starting dose	Dose level 1	Dose level 1
Cohort size	3	3
Sample size	Same as the median sample size obtained from TEQR design	Median sample size is automatically determined (not an input)
Number of dose levels	Same as the maximum dose levels examined (obtained from simulations) for the 3+3 design	Same as the maximum dose levels examined (obtained from simulations) for the 3+3 design
DLT probability deemed to be too toxic to allow further study at that dose level	NA	0.34
Desired sample size at MTD	NA	12
Maximum number of cohorts	NA	30
True DLT rate at each dose level	Values from Table 2 for each dose-toxicity curve	Values from Table 2 for each dose-toxicity curve

The mTPI software (R code) is available at:
http://health.bsd.uchicago.edu/yji/software2.htm
R code for the TEQR design was developed using the package TEQR.

Appendix Table 3: Parameters for the BOIN Design:

Parameter	BOIN Design
Number of simulations	2000
Target toxicity probability p_T	0.2
The interval for the target toxicity probability	Used the Default Interval Determined by the design, which is (0.16, 0.24) for p_T=0.2, and is very close to the interval (0.15, 0.25) used for the other model based designs.
Starting dose	Dose level 1
Cohort size	3
Sample size	Same as the median sample size obtained from TEQR design (the sample size is not a direct input of the program but the number of cohorts is an input and we input the number of cohorts such that the number of cohorts*cohort size is the desired sample size).
Number of cohorts	Desired sample size/cohort size
Cut off to eliminate an overly toxic dose for safety	0.95
True DLT rate at each dose level	Values from Table 2 for each dose-toxicity curve

R code for the BOIN design was developed using the package BOIN.

Appendix Table 4: Parameters for the CRM Design Used in CRMTrialSimulator:

Parameter	**CRM Design**
Number of simulations	2000
Max sample size	Same as the median sample size obtained from the TEQR design
Cohort size	3
Number of dose levels planned	Same as the maximum dose levels examined (obtained from simulations) for the 3+3 design
Starting dose	Dose level 1
Target toxicity probability	0.2
True DLT rate at each dose level	Values from Table 2 for each dose-toxicity curve
CRM Inputs:	
The probability of toxicity at dose i is modeled as $p_i^{\exp(\alpha)}$, where p_i is a constant and α is distributed a priori as a normal random variable	α is normally disturbed with mean 0 and variance 2
Prior probabilities of toxicity used are the defaults in the program	at dose level 1=0.15, at dose level 2=0.25, at dose level 3=0.3, at dose level 4=0.45, at dose level 5=0.51, at dose level 6=0.56, at dose level 7=0.6
Stopping probability (the trial is stopped if the probability that the lowest dose is more toxic than the target is greater than this value)	0.9

The software can be found at: https://biostatistics.mdanderson.org/SoftwareDownload/SingleSoftware.aspx?Software_Id=13

After the first cohort, each successive cohort is given the dose whose posterior probability of toxicity given the data collected thus far is closest to the target, subject to one additional requirement: one cannot skip over an untried dose. If the method would otherwise skip over an untried dose, the lowest untried dose is given instead.

Appendix Table 5: Parameters for the EWOC Design Used in Web-EWOC Simulator:

Parameter	**EWOC Design**
Number of simulations	2000
Sample size	Same as the median sample size obtained from the TEQR design
Cohort size	3
Number of dose levels planned	Same as the maximum dose levels examined (obtained from simulations) for the 3+3 design
Starting dose	Dose level 1
Target probability of dose limiting toxicity	0.2
Probability of exceeding target dose (α)	0.25
Variable α increment (resource to control the dose escalation rate in the beginning of the trial)	0.04
Minimum dose and Maximum dose	100 and 500 are the default values (the allowable range is 0 to 500) and the doses are equally spaced
True DLT rate at each dose level	Values from Table 2 for each dose-toxicity curve
Prior distribution	$\rho_0 \sim$ Uniform(0, 0.2) (the prior for ρ_0, the probability of DLT at the minimum dose, is Uniform(0, 0.2)) $\gamma \sim$ Uniform(100, 500) (the prior for the maximum tolerated dose γ is Uniform(100, 500))

The EWOC software is available at:
https://biostatistics.csmc.edu/ewoc/ewocWeb.php

Appendix Table 6: Simulation Results: Linear Dose-toxicity: DLT rate=min (-0.071197+0.000811966*dose, 1) – Target DLT rate=0.1 for the Model-Based Designs and Dose Level 2 is the True MTD:

Design	% of times that dose level 2 is selected as the MTD	% of times that doses below the MTD (dose level 1) are selected as the MTD	% of times that doses above the MTD (dose levels 3 and above) are selected as the MTD	Average Number of Dose Levels Examined	Std of Dose Levels Examined	Max Dose Levels Examined	Median Dose Levels Examined	Average Number of Patients per Trial	Median Number of Patients per Trial	Median Number of DLTs per Trial	Average Sample Size at MTD	Average % of pts dosed at MTD	Average % of pts under-dosed	Average % of pts over-dosed
3+3	26.85	7.75	65.21	3.9	1.01	7	4	14.75	15	3	3.6	24.75	20.91	54.34
mTPI	55.75	13.7	30.5	7		7	7	24 (max)	24 (max)		9.5	39.4	22.76	37.84
TEQR	50	15	31	7		7	7	22.81	24		8.8	38.58	23.63	37.79
BOIN	55.9	14.5	29.5	7		7	7	24 (max)	24 (max)	2.9 (mean)	9.6	39.83	26.97	33.2
CRM	57	24	20	7		7	7	24 (max)	24 (max)	2.3 (mean)	8.9	37.03	36.24	26.73
EWOC	43.35	4.15	52.5	7		7	7	24 (max)	24 (max)		8.1	33.93	17.87	48.19

The sum of columns 2 to 4 may add up to <100% because the remaining small percentage of times, no dose level is selected as the MTD.
The default interval for the target DLT rate in the R package is used for the BOIN design.

Appendix Table 7: Simulation Results: Linear Dose-toxicity: DLT rate=min (-0.071197+0.000811966*dose, 1) – Target DLT rate=0.33 for the Model-based Designs and Dose Level 4 is the True MTD:

Design	% of times that dose level 4 is selected as the MTD	% of times that doses below the MTD (dose levels 3 and below) are selected as the MTD	% of times that doses above the MTD (dose levels 5 and above) are selected as the MTD	Average Number of Dose Levels Examined	Std of Dose Levels Examined	Max Dose Levels Examined	Median Dose Levels Examined	Average Number of Patients per Trial	Median Number of Patients per Trial	Median Number of DLTs per Trial	Average Sample Size at MTD	Average % of pts dosed at MTD	Average % of pts under-dosed	Average % of pts over-dosed
3+3	22.77	72.09	4.95	3.9	1.01	7	4	14.75	15	3	2.8	19.14	71.76	9.1
mTPI	44.1	42.2	13.7	7		7	7	24 (max)	24 (max)		6.0	24.98	67.16	7.86
TEQR	31	65	4	7		7	7	24.57	24		6.1	24.75	65.73	9.52
BOIN	43.2	39.5	17.2	7		7	7	24 (max)	24 (max)	5.1 (mean)	6.0	25	65.83	9.17
CRM	53	32	15	7		7	7	24 (max)	24 (max)	5.7 (mean)	7.2	30	59.38	10.63
EWOC	48.15	11.05	40.8	7		7	7	24 (max)	24 (max)		8.0	33.44	43.27	23.29

The sum of columns 2 to 4 may add up to <100% because the remaining small percentage of times, no dose level is selected as the MTD.
The default interval for the target DLT rate in the R package is used for the BOIN design.

Appendix Table 8: Effect of the location of the starting dose relative to the true MTD on the accuracy of MTD selection for the 3+3 design for the three linear dose-toxicity curves with different offsets shown in Appendix Figure 4:

	Starting dose level relative to true MTD				
Background DLT Rate	-6	-5	-4	-3	-2
.1	21.27%	23.91%	26.43%	29.24%	32.36%
.05	26.41%	28.22%	28.05%	29.45%	29.69%
0	29.86%	31.07%	30.44%	30.9%	30.38%

-6 implies that the starting dose is 6 dose levels below the true MTD, and similarly for the others. We observe that for an offset of 0 (when the true DLT rate=0 for the first 6 dose levels), the accuracy of MTD selection is not affected by how many dose levels below the true MTD the starting dose level is located i.e. the percentage of times (out of 10000 simulations) that dose level 6 (true MTD) is selected as the MTD is constant (~30%) for the different starting dose locations relative to the true MTD. However for an offset of 0.1 (when the true DLT rate=0.1 for the first 6 dose levels), the accuracy of MTD selection is affected by how many dose levels below the true MTD the starting dose level is located.

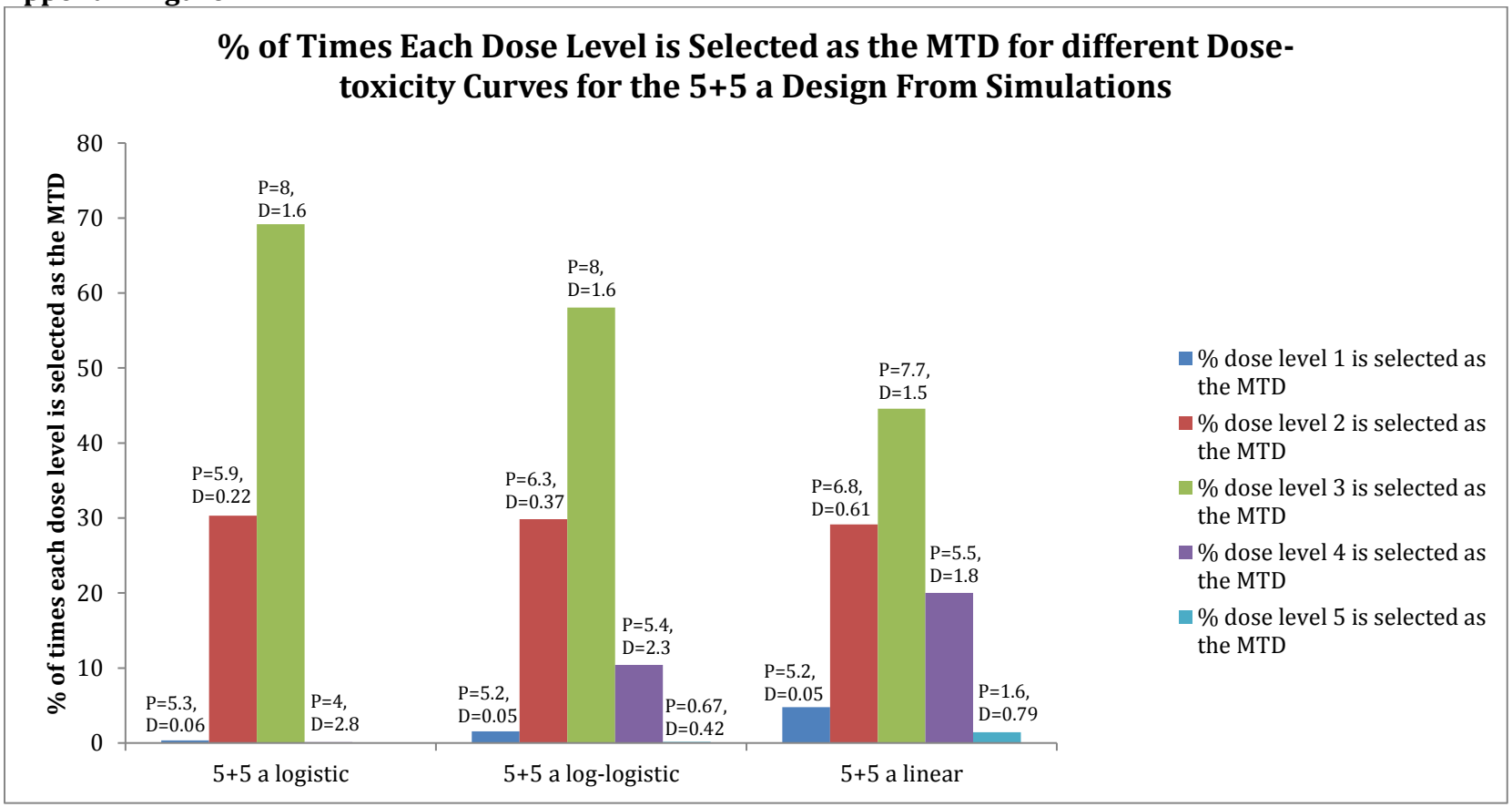
Appendix Table 9: Accuracy of MTD selection in the 3+3 design when the true DLT rate at one of the dose levels falls within the target DLT rate interval (0.17–0.26) and when more than one of the true DLT rates fall within the target DLT rate interval

Dose Level	DLT Rate	Accuracy of MTD Selection	DLT Rate	Accuracy of MTD Selection	DLT Rate	Accuracy of MTD Selection	DLT Rate	Accuracy of MTD Selection
1	0.01	1.5%	0.01	23.01%	0.01	25.09%	0.01	1.58%
2	0.04	28.28%	0.17	32.06%	0.18	46.46%	0.04	41.33%
3	0.2	68.05%	0.26	43.63%	0.36	27.60%	0.26	55.23%
4	0.71	2%	0.71	1.16%	0.71	0.7%	0.71	1.68%
5	0.97	0%	0.97	0%	0.97	0%	0.97	0%

This table provides an example of when the exact target DLT rate is not one of the true DLT rates.
The second column contains the true DLT rates based on the logistic dose-toxicity curve in Table 2. The 3+3 design targets a DLT rate of ~0.2 with an interval of (0.17, 0.26). For the columns where the true DLT rate at exactly one of the dose

levels falls within this interval (i.e. columns 2, 6 and 8), that dose level is the true MTD. However, for column 4, the true DLT rates at two of the dose levels fall within this interval, and both these dose levels can be considered the true MTD. In this case, the accuracy of MTD selection is 75.69% (=32.06%+43.63%). For columns 3, 7 and 9, the accuracy of MTD selection is highlighted in blue. Note also that the sum of each column in columns 3, 5, 7 and 9 may add up to <100% because the remaining small percentage of times, no dose level is selected as the MTD. Also see Appendix Figure 6.

Appendix Figure 1:

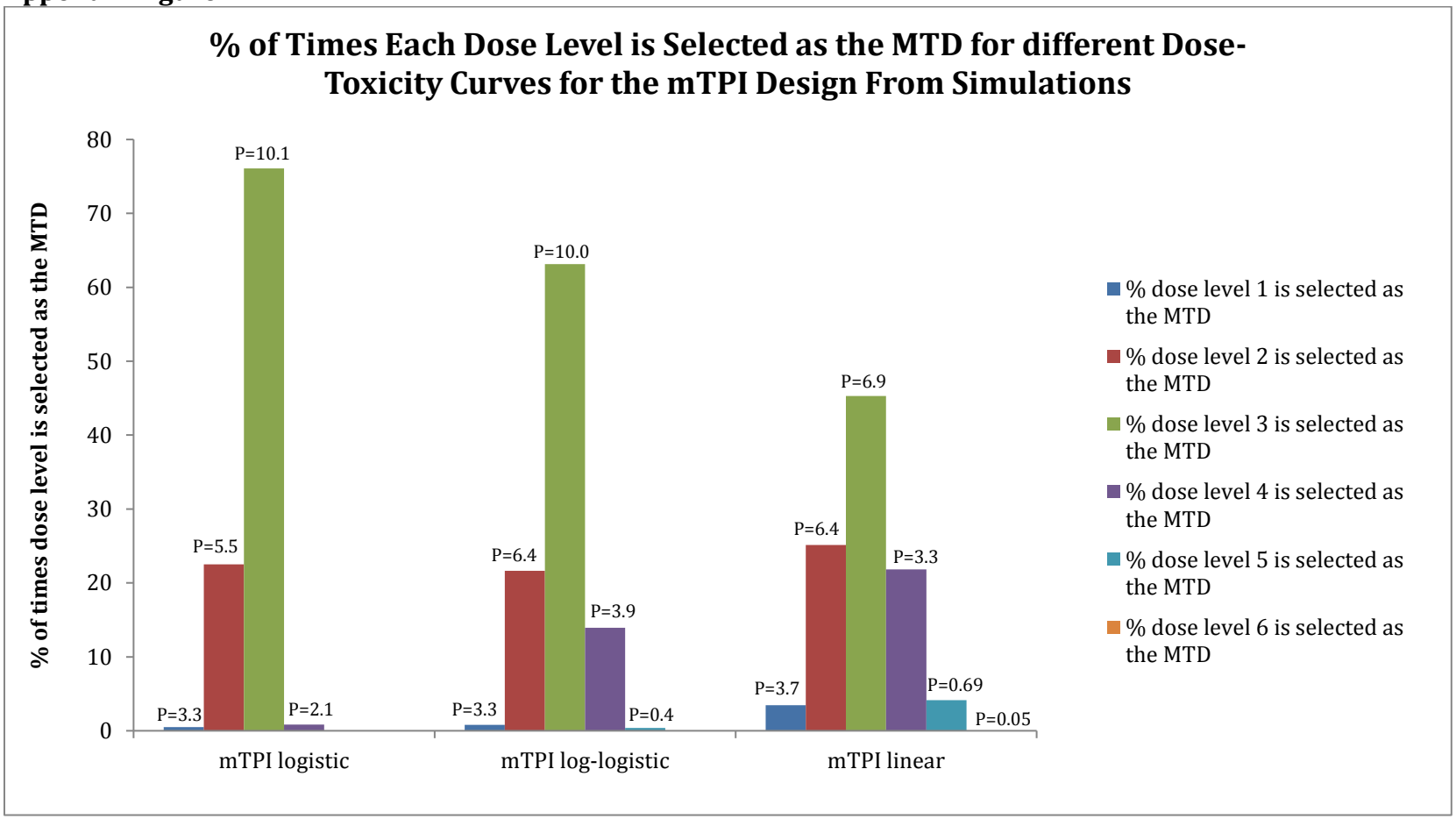


5+5 a logistic implies the 5+5 a design with the true DLT rates given in Table 2, generated from the logistic dose-toxicity curve, and similarly for the others.

P=Mean Sample Size and D=Mean Number of DLTs at each dose level (from 10000 simulations).

The figure depicts the percentage of times that the 5+5 a design selects each dose level as the MTD for the true DLT rates given in Table 2, generated from the three dose-toxicity curves. These percentages are from simulations and the results are shown in Tables 3–5.

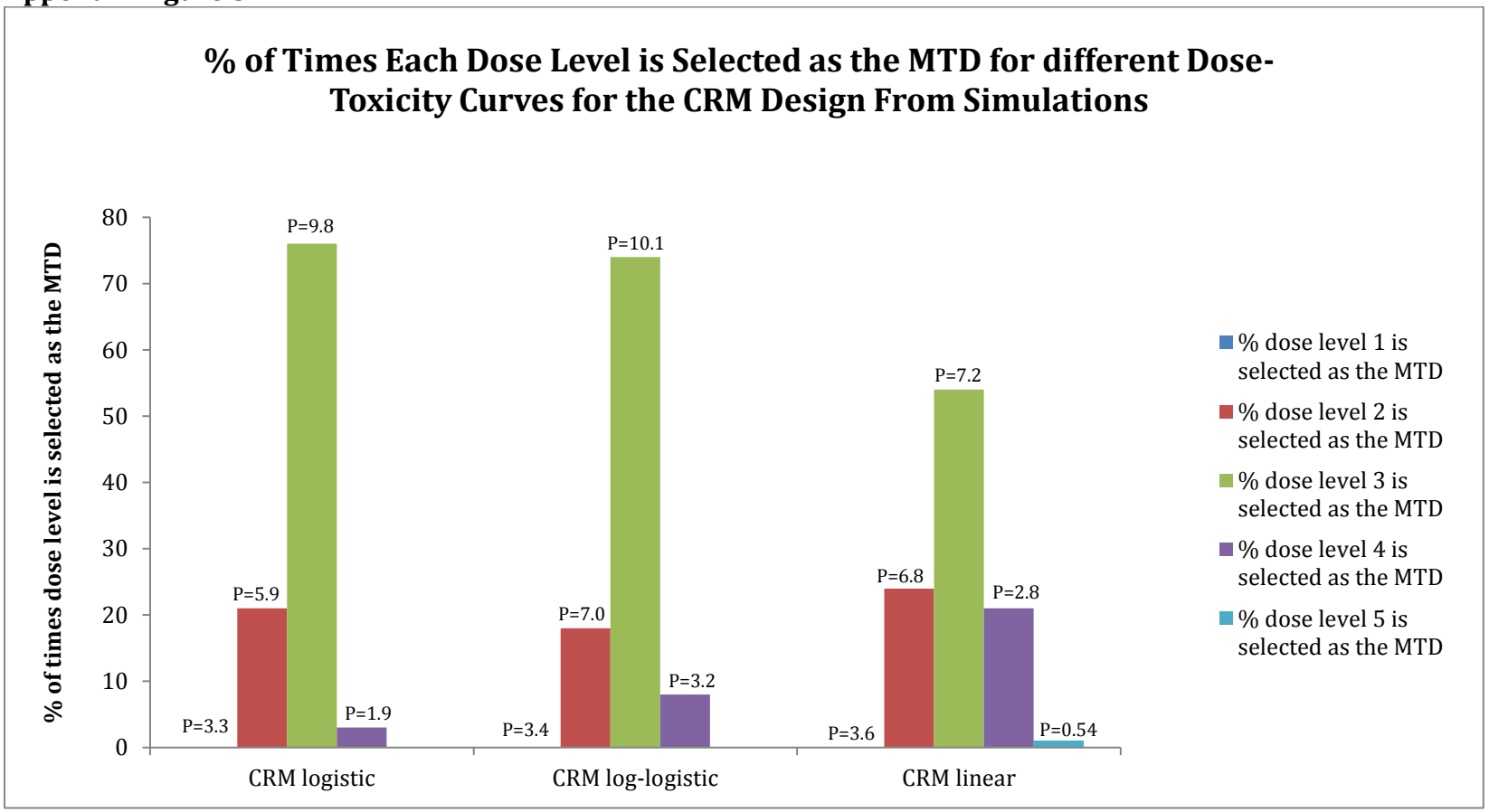
Appendix Figure 2:



mTPI logistic implies the mTPI design with the true DLT rates given in Table 2, generated from the logistic dose-toxicity curve, and similarly for the others. P=Mean Sample Size at each dose level (from 10000 simulations).

The figure depicts the percentage of times that the mTPI design selects each dose level as the MTD for the true DLT rates given in Table 2, generated from the three dose-toxicity curves. These percentages are from simulations and the results are shown in Tables 3–5.

Appendix Figure 3:

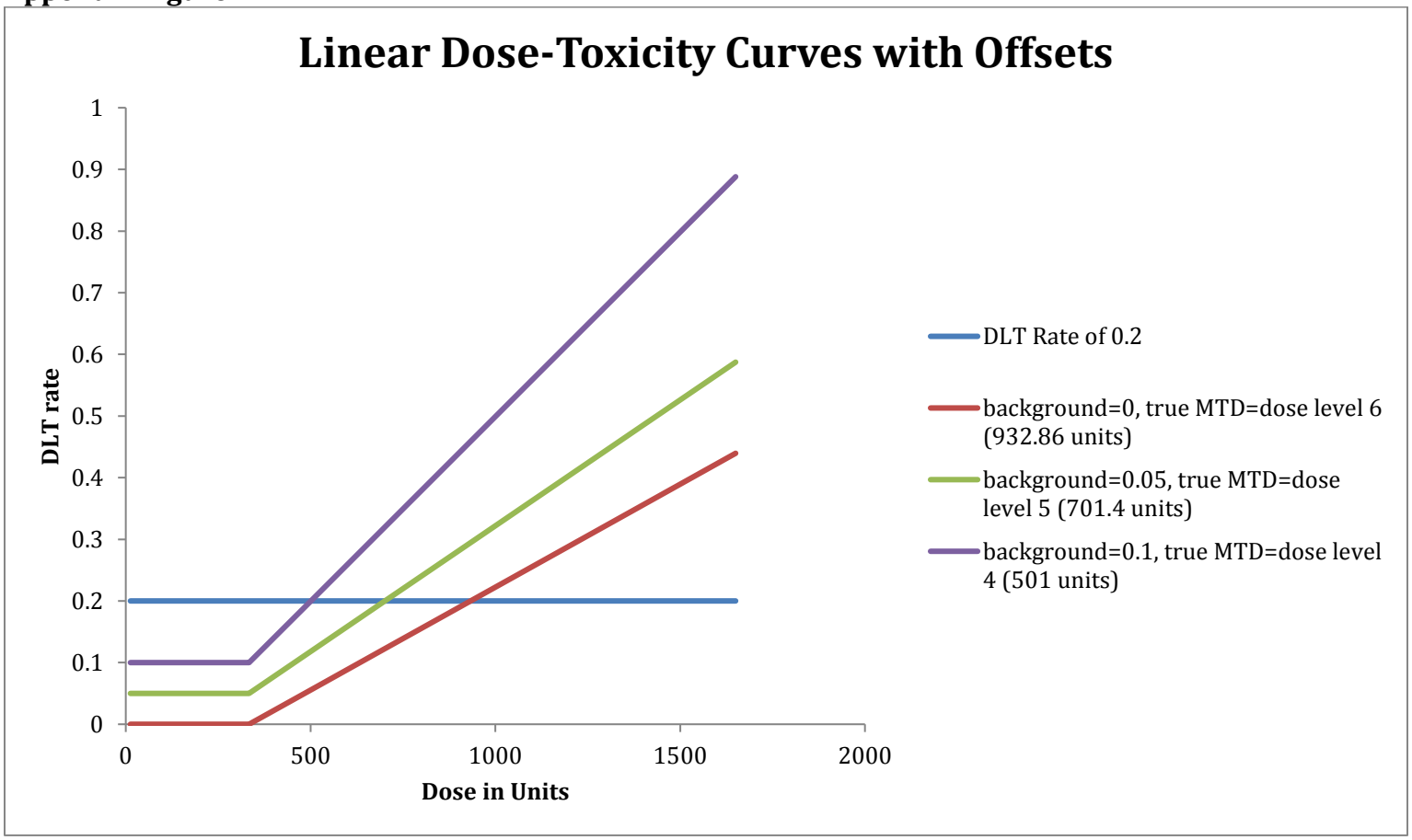


CRM logistic implies the CRM design with the true DLT rates given in Table 2, generated from the logistic dose-toxicity curve, and similarly for the others.

P=Mean Sample Size at each dose level (from 10000 simulations).

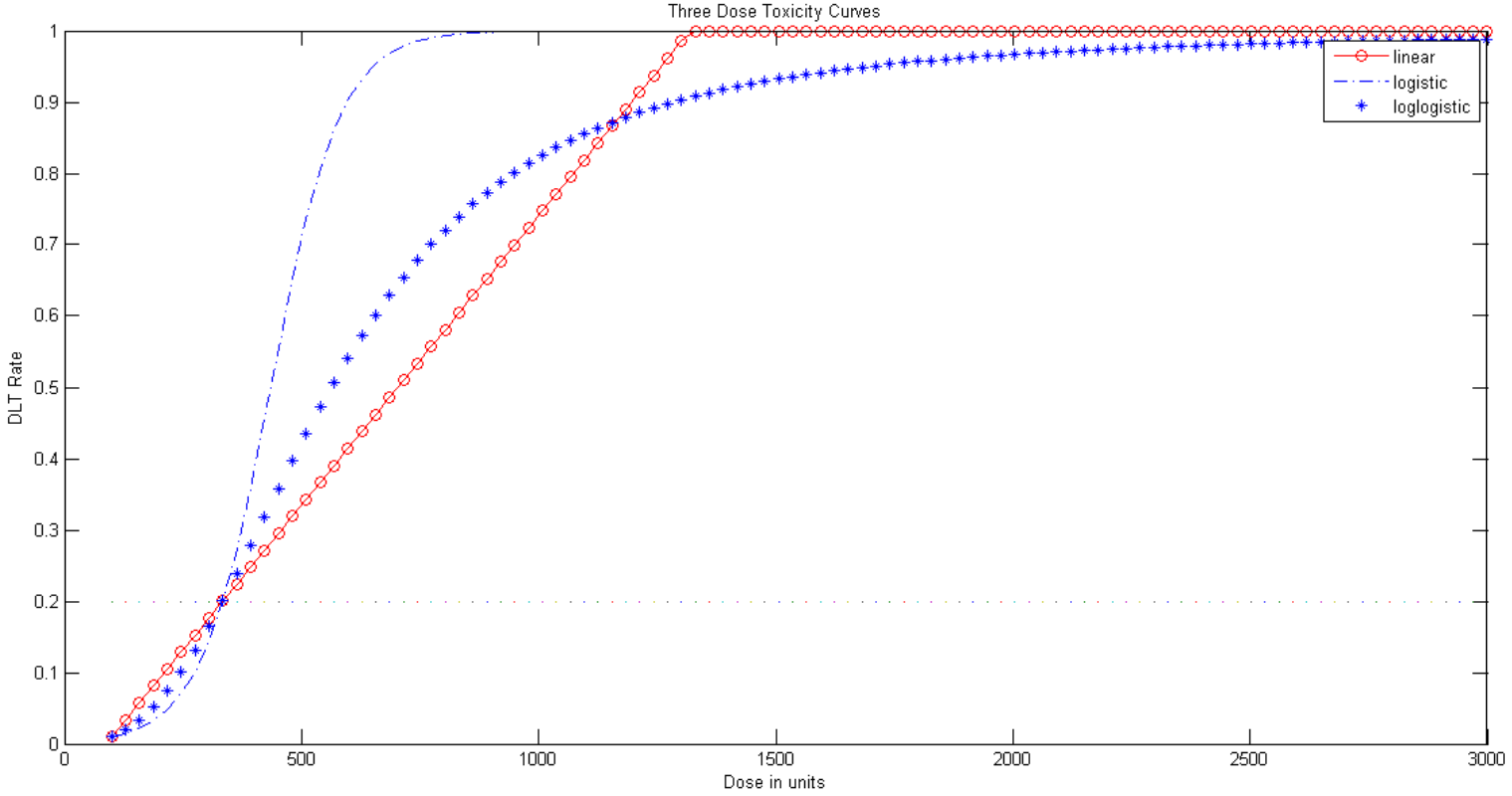
The figure depicts the percentage of times that the CRM design selects each dose level as the MTD for the true DLT rates given in Table 2, generated from the three dose-toxicity curves. These percentages are from simulations and the results are shown in Tables 3–5.

Appendix Figure 4:



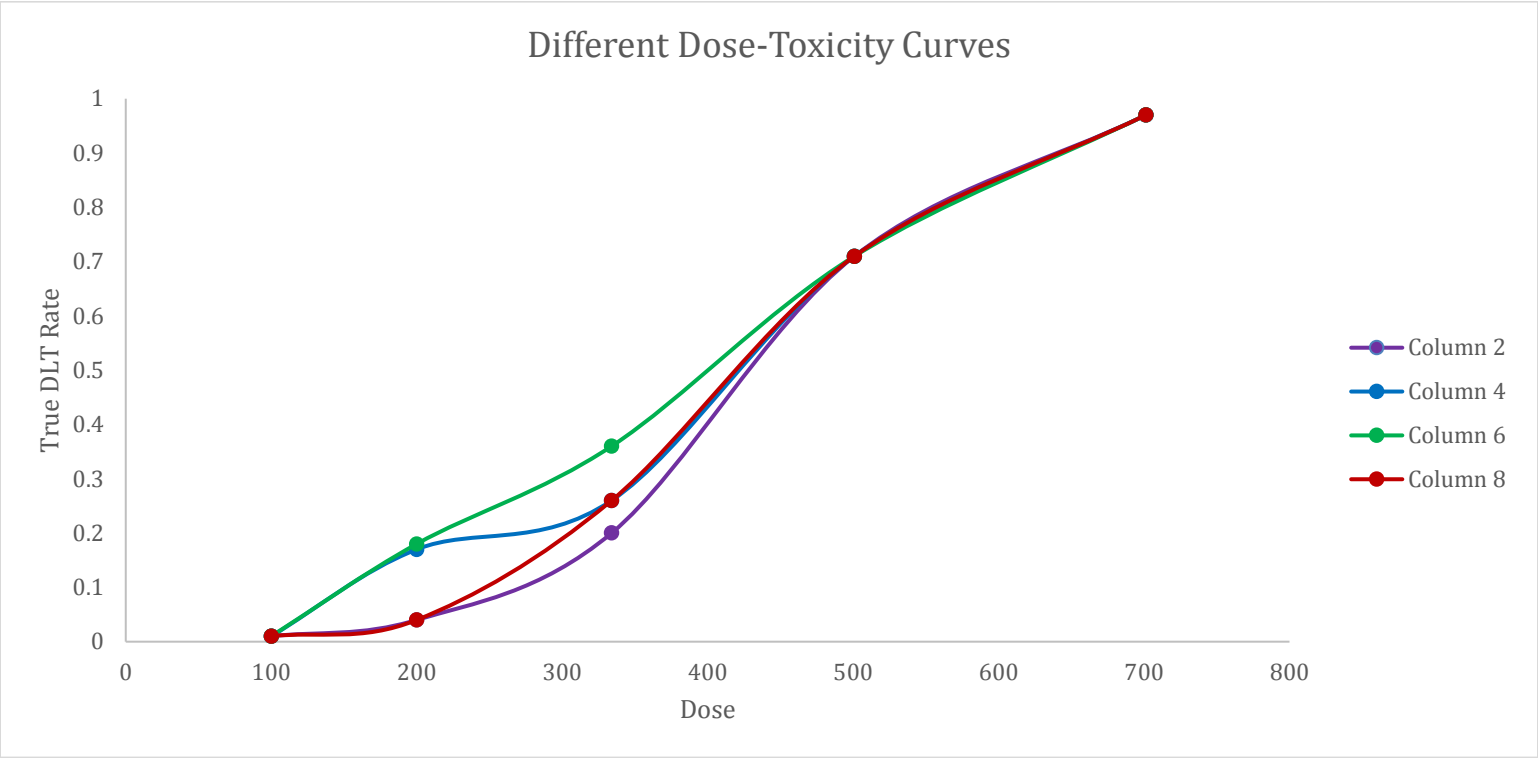
The figure depicts three linear dose-toxicity curves with different offsets that are used to investigate the effect of the location of the starting dose relative to the true MTD on the accuracy of MTD selection for the 3+3 design.

Appendix Figure 5:



This plot depicts the dose-toxicity curves that correspond to the true DLT rates in Table 2. The separation between the true DLT rate at the MTD and that at the level below and above it is large for the logistic dose-toxicity curve, but is much smaller for the linear dose-toxicity curve. For the log-logistic dose-toxicity curve, the separation between the true DLT rate at the MTD and that at the level below and above it is in between that of the logistic and linear dose-toxicity curves.

Appendix Figure 6:



This figure depicts the dose-toxicity curves that correspond to the true DLT rates in Columns 2, 4, 6 and 8 of Appendix Table 9.

CHAPTER THREE

Introduction

In a Phase 1 oncology clinical trial, the safety of the investigational drug is studied and the maximum tolerated dose (MTD) of the study drug to be used in a Phase 2 trial is determined [91, 77]. In a Phase 2 trial, in addition to safety, the efficacy of the drug is investigated. The number of patients enrolled in a Phase 1 trial is usually very small (say 15–30 patients), while the number of patients enrolled in a Phase 2 trial is larger (say 50–100).

With such a small number of patients, Phase 1 dose-finding oncology trials do not accurately select the MTD. The 3+3 design that is still frequently used in many Phase 1 dose-finding oncology trials has been shown to be inaccurate in determining the MTD and to under-dose a large number of patients [68]. Furthermore, the efficacy of an anti-cancer agent may not always increase with an increase in dose, and can peak at any dose level [92,93]. Thus, Phase 1 trials may not target the optimal dose taking into consideration both toxicity and response. Assessment of both dose limiting toxicities (DLTs) and efficacy responses in a reasonably sized, larger trial has the potential for a more accurate determination of a suitable dose, compared to a Phase 1 followed by a Phase 2 trial.

The criticality of selecting the right safe dose and the optimal dose for toxicity and

efficacy has been illustrated in several cases. In the paper by Markman [94], the drug pegylated liposomal Doxorubicin (PLD), which is used to treat platinum resistant ovarian cancer, and its dosage are discussed. The dose of PLD explored in initial clinical trials as well as the dose approved by the FDA was 50 mg/m^2 (every 28 days). However clinical experience has clearly shown that this dose leads to substantial adverse effects in patients, while a lower dose of 40 mg/m^2 is equally efficacious but leads to fewer adverse effects, as emphasized by Ferrandina [95]. As other examples, the reference by Schilsky [96] provides a summary in Table 1 of some approved oncology drugs along with their approved dosages and states that for many of these drugs, the dose that is clinically administered is often lower or higher than the dosage that was approved by the regulatory authority. The dose that is commonly used in a Phase III oncology trial is the MTD determined from a small number of patients treated in an earlier dose-finding Phase I trial, as Phase II oncology trials do not typically evaluate multiple doses. If the MTD obtained from the Phase I trial is not right, this can have enormous cost and resource repercussions for the development of an oncology drug. If the dose administered to patients in the Phase III trial is too high, the trial can fail because of many early drop-outs due to adverse events. On the other hand, If the dose administered to patients is inefficacious, it may not be possible to determine from the Phase III trial results whether the drug does not work at all or whether just the dosage was wrong. All this may result in having to conduct another new Phase III trial with the optimal dose. Thus, conducting small early phase trials and rapidly moving on to a Phase III trial may not be a good strategy in the long run,

and it may be worthwhile to spend more time in the early phase trials to thoroughly investigate and accurately determine the optimal drug dose for safety and efficacy.

Hence, in this paper, we propose a simple integrated Phase 1/2 trial design called the "*20+20 accelerated titration design*" that evaluates both safety and efficacy with a somewhat larger sample size. The motivation for exploring a larger rule-based design comes from the substantial effect of sample size and/or cohort size on the accuracy of dose selection in various early phase oncology designs. With 35–50 patients, the safety of the drug and the MTD can be evaluated with greater accuracy than with 15–30 patients, and the efficacy of the study drug can also be investigated in this larger sample. Thus, a single, larger study could serve to determine a dose that is optimal for both safety and efficacy.

Designs such as 1) the seamless Phase 1/2 SEARS design [97, 98], 2) a seamless 2-step Phase 1/2 design [99, 100], 3) designs to find the optimal biological dose [93, 101] and 4) the Eff-Tox design [67, 102] among others [103, 104, 105, 106] have been proposed to evaluate both drug toxicity and efficacy in early phase studies; a recent reference by Yuan covers Bayesian designs for Phase I/II trials in detail [107]. In fact, our proposed rule-based 20+20 accelerated titration design is similar in concept to the Eff-Tox design. The Bayesian decision rules incorporated in our design at the end of the trial for selecting an optimal dose for efficacy and safety are the same as those used in the Eff-Tox design for determining the dose level to which the next cohort of

patients should be assigned, based on the number of DLTs and responses in previous patients. However, unlike these other designs, our rule based design is easily implemented without the assistance of a statistician during the conduct of the trial. In addition, our design does not require the challenging selection of three points to define the trade-off function contour used to determine the optimal dose for the next cohort of patients in the Eff-Tox design. In the Eff-Tox design, if the trade-off contour is not sufficiently steep, the design can get stuck at a lower dose and then fail to find an optimal dose for the assignment of the next cohort of patients (Eff-Tox tutorial[5]).

In summary, we first systematically study the role of sample size and cohort size on the accuracy of dose selection in the Continual Reassessment Method (CRM), Eff-Tox and some rule-based designs. We then use this information to propose a design that is large enough to allow accurate dose selection for toxicity and that selects an optimal dose for both toxicity and efficacy via Bayesian decision rules at the end. In particular, we propose the *20+20 accelerated titration* design. We use simulations to study the statistical operating characteristics of the 20+20 accelerated titration design and compare its performance with that of the Eff-Tox design by Thall et al. [67, 102] and the Optimal Biological Dose (OBD) Isotonic design by Zang et al. [101]. We show via simulated examples that it performs as well as or better than the Eff-Tox

5 https://biostatistics.mdanderson.org/softwaredownload/ProductSupportFiles/EffTox/EffToxUsersGuide.pdf

design and the OBD Isotonic design in these situations. We also demonstrate that the proposed design performs better than a 3+3 Phase 1 design followed by a Phase 2 design.

Methods

Effect of Sample Size/Cohort Size on the Accuracy of MTD or Optimal Dose Selection

We first study via simulations the effect of sample and/or cohort sizes on the accuracy of MTD selection in the CRM as well as the rule based designs such as the 3+3, 5+5 a, 10+10 and 20+20 designs, using a logistic dose-toxicity curve, which is often employed to describe the relation between dose and toxicity [84]; we also study the effect of sample size and cohort size on the accuracy of optimal dose selection in the Eff-Tox design [67] [see Simulation Section for details]. We describe each design studied in this paper in brief below.

Design Descriptions

The CRM design is a Bayesian design, where the cumulative DLT data along with a pre-specified dose-toxicity model, frequently a one or two parameter logistic model, are used to assign the next patient(s) to a dose level [66, 84]. After the DLT evaluation period of each patient (or cohort of patients), the parameters of the dose-toxicity model are updated. The patient(s) is assigned to the dose level whose DLT rate is closest to but less than the estimated DLT rate from the updated model. The stopping

point of the trial is usually the pre-specified sample size.

The Eff-Tox design is also a Bayesian design. It considers the trade-off between the probabilities of drug toxicity and efficacy to determine the optimal dose for each new cohort of patients. The joint probability π_{ab} of efficacy and toxicity at each dose is calculated in terms of the marginal probabilities of efficacy and toxicity (π_E and π_T). The equation for π_{ab} in terms of π_E and π_T and its role in computing the dose level with maximum desirability based on the cumulative DLT and efficacy data are detailed in Thall et al. [102]. Before each new cohort of patients is enrolled, the desirability of each dose level is calculated; the dose level with the maximum desirability is the dose level to which the next cohort of patients is assigned. The stopping point of the trial is usually the pre-specified sample size.

In the paper by Zang et al. [101], three designs are proposed to estimate the optimal biological dose (OBD), namely the OBD logistic design, the OBD Isotonic design and the OBD locally-logistic design. The authors recommend the OBD Isotonic design or the OBD locally-logistic design for use in practice, based on their robust operating characteristics from simulation studies. We focus only on the OBD Isotonic design in this paper due to the reasonable run-time for 10000 simulations for the OBD Isotonic design compared to that for the OBD locally-logistic design. The details of the OBD Isotonic design are provided in the reference by Zang et al.. To determine the OBD, an admissible set, which is the set of doses satisfying a safety criterion similar to that

used in the Eff-Tox design, is first defined. The OBD is then the lowest dose with the highest response rate within the admissible set of doses, while still being safe. The stopping point of the trial is usually the pre-specified sample size.

The 3+3 and 5+5a rule based designs target a DLT rate Γ of $\sim$0.2 (see Table 4.1 of Chapter 4 of the reference by Ting, 2006 [89] and Table 1 of Ananthakrishnan, 2017 [108]).

3+3 Design: Enroll 3 patients at the lowest dose level.

- If 0 out of 3 enrolled patients have a DLT, then escalate to the next dose level and enroll 3 more.
- If 1 out of 3 patients has a DLT, then add 3 more patients at the same dose level; if 2 or more patients out of 3 or 6 patients experience a DLT, then stop the trial.
- The MTD is one dose level below the last dose level examined.

5+5 a Design: Enroll 5 patients at the lowest dose level.

- If 0 out of 5 enrolled patients have a DLT, then escalate to the next dose level and enroll 5 more.
- If 1 or 2 out of 5 patients have a DLT, then add 5 more patients at the same dose level; if 3 or more patients out of 5 or 10 patients experience a DLT, then stop the trial.
- The MTD is one dose level below the last dose level examined.

The 10+10 and 20+20 designs we construct are such that they target a DLT rate Γ of ~0.2, (see the Appendix for the target DLT interval of the 20+20 design).

10+10 design: Enroll 10 patients at the lowest dose level.

- If <=2 patients out of 10 experience a DLT, then enroll 10 patents in the next higher dose level.
- If 3 or 4 patients experience a DLT, then enroll 10 more patients in the same dose level.
- If <=4 patients out of 20 experience a DLT, then enroll 10 patients in the next higher dose level.
- If 5 or more patients out of 10 or 20 patients experience a DLT at a dose level, then stop the trial.
- The MTD is one dose level below the last dose level examined.

20+20 design: Enroll 20 patients at the lowest dose level.

- If <=6 patients out of 20 experience a DLT, then enroll 20 patents in the next higher dose level.
- If 7 or 8 patients experience a DLT, then enroll 20 more patients in the same dose level.
- If <=8 patients out of 40 experience a DLT, then enroll 20 patients in the next higher dose level.
- If 9 or more patients out of 20 or 40 patients experience a DLT at a dose level, then stop the trial.

- The MTD is one dose level below the last dose level examined.

Proposed Design

We propose the *20+20 accelerated titration* design that is intended to be a simple rule based design during the conduct of the study, and which enrolls larger sample sizes than standard Phase 1 designs and incorporates Bayesian decision rules for optimal dose selection at the end of the trial. As mentioned, the 20+20 design targets a DLT rate of ~0.2, similar to the 3+3 design, but our proposed 20+20 accelerated titration design has additional safety stopping rules which decrease the target DLT rate of the 20+20 part of the design.

The schematic of the 20+20 accelerated titration design is shown in Figure 8 and the design is described here. The 20+20 accelerated titration design starts in the accelerated titration phase by enrolling patients in cohorts of size 3.

- The first cohort of 3 patients is assigned to the lowest dose level and subsequent cohorts of size 3 continue to be assigned to increasing dose levels as long as none of the 3 patient cohorts experience a DLT.
- When one or more patients in a cohort of size 3 experiences a DLT, then the design switches from the accelerated titration phase into the 20+20 design phase, where 20 patient cohorts are enrolled in batches of 6 or 8 patients (To limit the number of patients that are exposed to the study drug at once, we implement stopping rules

after every 6 or 8 patients i.e. after patient 6, 14, 20 etc.).

In other words, in our proposed design, 1 or more DLTs in a cohort of 3 patients leads to enrolling 3 new patients at that same dose level for an initial total of 6 patients.

- If 4 or more out of these first 6 patients have a DLT, then the trial is stopped.
- If <4 patients out of these first 6 patients have a DLT, then 8 more patients are treated at the same dose level.
    - If out of the 14 patients, 9 or more patients have a DLT, then the trial is stopped.
    - If out of the 14 patients, 8 or less patients have a DLT, then 6 more patients are treated at the same dose level.
        - If out of the 20 patients, 6 or less patients have a DLT, then initially 6 patients are treated at the next higher dose level and the process is continued, but if out of the 20 patients, 9 or more have a DLT, then the trial is stopped.
        - If out of the 20 patients, 7 or 8 patients have a DLT, then the same dose level is expanded, and 6 more patients are treated at the same dose level.
            - If out of 26 patients, more than 8 patients have a DLT, then the trial is stopped.
            - If out of 26 patients, no more than 8 patients have a DLT, then 8 more patients are treated at the same

dose level.

- If out of 34 patients, more than 8 patients have a DLT, then the trial is stopped.
- If out of 34 patients, no more than 8 patients have a DLT, then 6 more patients are treated at the same dose level.
    - If out of 40 patients, more than 8 patients have a DLT, then the trial is stopped.
    - If out of 40 patients, no more than 8 patients have a DLT, then initially 6 patients are treated at the next higher dose level and the process is continued.

Thus, we can escalate to the next dose level if 6 or less out of 20 patients in a dose level experience a DLT or 8 or less out of 40 patients in a dose level experience a DLT, similar to the escalation rules of the 20+20 design. However, we can also escalate if out of the first 14 patients in a dose level, we observe 0 DLTs and 0 responses, since no more than 6 DLTs can be observed in the last 6 patients and the dose is not efficacious. Also, note that a safety committee that can stop the trial at its discretion at any point should be implemented.

At the end of the trial, a dose that is acceptable for safety and efficacy is chosen using the following Bayesian decision rules on the posterior probabilities of p_i and q_i, which are the toxicity and efficacy probabilities at dose i:

$\Pr(p_i<p_{TP}|data) > a$ and $\Pr(q_i>p_{EP}|data) >b$

p_{TP} and p_{EP} are the upper limit for toxicity and lower limit for efficacy respectively, whose values are pre-specified for the study based on discussions with the clinician, and "a" and "b" are small probability cut-offs[6]. In this paper, we use $p_{TP}=0.33$, $p_{EP}=0.5$ and $a=b=0.1$. We also assume that both p_i and q_i follow a Jeffreys prior Beta(0.5, 0.5), which is a uniform prior from 0 to 1 [98]. The posterior distributions of p_i and q_i are then Beta($0.5+x_i$, $0.5+n_i-x_i$) where x_i is the number of DLTs or responses respectively at dose level i out of n_i patients at that dose level. These Bayesian decision rules are the same as those used in the Eff-Tox design to assign new cohorts of patients a dose, and they ensure that doses that are too toxic or that are too inefficacious are not selected.

If more than one dose is found to be acceptable for safety and efficacy in a trial using the Bayesian decision rules above, then the following criteria can be used, in the suggested order of preference, to choose a single dose level: a) the value of a pre-specified utility function evaluated at each dose level or b) the percentage of patients

[6] The cutoff probabilities are typically 0.1 or smaller in value (https://biostatistics.mdanderson.org/softwaredownload/ProductSupportFiles/EffTox/EffToxUsersGuide.pdf), and we use the upper limit of 0.1 for all the simulations in this paper.

who respond but do not have a DLT at each dose level or c) the empirical odds ratio (OR)[7] at each dose level. In this case, we would select the dose that has the maximum value for the utility function or that has the largest percentage of patients with a response but no DLT or that has the smallest value for the empirical odds ratio. We provide an example utility function for criterion a) and show the calculation for the empirical OR for criterion c) below.

One possible utility function would be the fraction of responders (out of the total patients) in the dose level minus a constant 'c' times the fraction of patients with DLTs. Other more complex utility functions that consider additional factors such as trial cost could be used as well, but here we use the following simple utility function. So, for dose level i, the following formula is used:

$utility_i = r_i - c*d_i$, where c is a constant that can vary between 0 and 1, r_i is the fraction of patients having a response, and d_i is the fraction of patients with a DLT. This utility function was also employed in Ivanova et al. [109].

The calculation for the odds ratio is based on the numbers of subjects, the number of responses, and the number of DLTs at a dose level. The following formula is used:

$$OR_i = \frac{\text{number of DLTs at dose i}*(\text{number of patients at dose i} - \text{number of responses at dose i})}{\text{number of responses at dose i}*(\text{number of patients at dose i} - \text{number of DLTs at dose i})}$$

[7] http://www2.ims.nus.edu.sg/Programs/011wclinic/files/guosheng_ppt.pdf

Design Characteristics and Comparisons

We then compare the results for accuracy of optimal dose selection of the 20+20 accelerated titration design to those of the Eff-Tox design and the OBD Isotonic design for various scenarios of true toxicity and efficacy rates. We also compare the results of the 20+20 accelerated titration design to those of a 3+3 Phase 1 design followed by a Phase 2 design.

Simulations

1) Effect of sample size and cohort size on the accuracy of MTD or dose selection

a) For the rule-based designs, we use the statistical package *SAS* to simulate and study the effect of sample size on the accuracy of MTD selection, as described in Ananthakrishnan et al. [108].

b) For the CRM design, we use the *R* package *CRM* to study the effect of sample size and cohort size on the accuracy of MTD selection with the input parameters given in the Appendix.

c) For the Eff-Tox design, we use the *Eff-Tox design* package from the MD Anderson Cancer Center to study the effect of sample size and cohort size on the accuracy of dose selection with the input parameters given in the Appendix.

2) 20+20 accelerated titration design

We create our own SAS program to simulate the proposed 20+20 accelerated titration design. We start by generating two correlated binary random variables X1~

Bernoulli(p) and X2~ Bernoulli(q) for toxicity and efficacy respectively as follows. We first generate X1 with success probability p. If X1 = 1, we generate X2 with probability q1, and if X1=0, we generate X2 with probability q2. r is the correlation coefficient between X1 and X2. Here,

q1 = q + (r/p) * sqrt(p*(1-p)*q*(1-q))

q2 = (q - q1*p)/(1-p)

The correlation coefficient r is restricted and can lie only between

$\max(-(\frac{(pq)}{(1-p)(1-q)})^{\frac{1}{2}},\ -(\frac{(1-p)(1-q)}{(pq)})^{\frac{1}{2}})$ and $\min((\frac{(p\,(1-q))}{(1-p)(q)})^{\frac{1}{2}},\ (\frac{((1-p)q)}{(p)(1-q)})^{\frac{1}{2}})$ [110].

From the above equations, it can be seen that if the correlation coefficient r=0 then q1=q2=q.

In simulations to study the statistical operating characteristics of the 20+20 accelerated titration design, we generate the true DLT rate at each dose level (p_i) using a logistic dose-toxicity curve because the toxicity of an anti-cancer agent typically increases with an increase in dose; the two coefficients of the logistic dose toxicity curve are calculated using the parameters in Table 7. However, we select the true response rate at each dose level (q_i) manually because the efficacy of an anti-cancer agent may not always increase with an increase in dose, and can peak at any dose level [92, 93]. In Table 8, we select the true response rates such that the true response rate peaks at dose level 4. However, different dose-response curves can be investigated (see results in the Appendix for the 20+20 accelerated titration design for other scenarios of true underlying response rates).

Using SAS, we simulate the 20+20 accelerated titration design proposed 10000 times along with the true DLT and response rates in Table 8, correlation coefficient r=0, and the stopping rules described in the Methods section. We start with a correlation coefficient r of 0 in our simulations for simplicity, but calculations with other non-zero correlation coefficients are discussed briefly in the Results Section and the Appendix. For each simulated trial, we estimate the number of patients, the number of DLTs and the number of responses in each dose level. We then use the Bayesian decision rules described, to select an acceptable dose for toxicity and efficacy at the end of each simulated trial. We also estimate the value of the utility function at each dose, the percentage of patients at each dose who respond but do not have a DLT and the empirical odds ratio at each dose, for use in optimal dose selection.

Results

Effect of Sample and Cohort Size on Dose Selection

a) Phase 1 designs (safety)

Table 9 shows the percentage of times that the true MTD (dose level 4) is selected by various rule based designs that allow only escalation and that target a DLT rate of ~0.2. These results are for the true DLT rates shown in Table 8 and are based on 10000 simulations for the rule-based designs. The accuracy of MTD selection, the median and maximum sample sizes are outputs of the simulations for these rule-based designs. For comparison, we also include in Table 9 the results of two specific cases of the CRM design using the input parameters given in the Appendix

(further scenarios for the CRM design are shown in Table 10). As can be seen, the accuracy of MTD selection increases with an increase in sample size for all the cases considered. For the 20+20 design, the accuracy of MTD selection is very high (~90%), but the median number of patients enrolled in the trial is also high (100).

All the results in Table 10 for the CRM design are based on 2000 simulations for the input parameters given in the Appendix. The sample size and cohort size are inputs and the accuracy of MTD selection is an output of these simulations. As in Table 9, the percentage of times that the true MTD (dose level 4) is selected increases with an increase in sample size. For a sample size of 50 and a cohort size of 5 for example, the true MTD (dose level 4) is selected ~82% of the time. It is also observed that the cohort size, given the same total sample size, does not have a large effect on the accuracy of MTD selection in the CRM design, if the cohort size is a small percentage of the total sample size.

b) Eff-Tox design (safety and efficacy)

The effect of sample size and cohort size on dose selection in the Eff-Tox design can be seen using the example in Table 11, which is based on an example in the Eff-Tox website https://biostatistics.mdanderson.org/softwaredownload/ProductSupportFiles/EffTox/EfftoxTutorial.html. All the results in Table 11 are based on 10000 simulations using the input parameters given in the Appendix for the Eff-Tox design. We expect

dose level 5, the dose level with the highest trade-off value between the probabilities of efficacy and toxicity in Table 11, to be selected most frequently. However, we observe that when the sample size is small (18) and the cohort size is 3, dose level 4 is selected more frequently than dose level 5. With an increase in sample size to 99 and the same cohort size of 3, dose levels 4 and 5 are selected with equal frequency. With a sample size of 99 but a cohort size of 9, dose level 5 is selected more than twice as frequently as dose level 4. Hence, cohort size appears to be an important criterion in dose selection in the Eff-Tox design.

In summary, based on these examples of the rule-based, CRM and Eff-Tox designs, the accuracy of MTD or dose selection improves dramatically with an increase in sample size for all the cases and designs considered. Larger cohort sizes may result in a small reduction in the accuracy of MTD selection for the CRM design, but could improve dose selection in the Eff-Tox design. Thus, cohort size and sample size are crucial parameters to consider and explore, while designing an early phase oncology trial.

Simulation Results for the 20+20 Accelerated Titration Design

From the simulation results in the previous section, we observe that the 20+20 design has a high probability of selecting the MTD accurately due to its larger sample size than the other A+B designs. Our simulations yield a median sample size of 100 for the 20+20 design for the true DLT rates in Table 8. This larger sample size allows us to consider addressing drug efficacy in addition to drug safety. However, the sample size

of the 20+20 design is relatively large for an integrated Phase 1/2 trial, and a sample size closer to 50 would be more reasonable. Therefore, we propose the 20+20 accelerated titration design, as our simulations yield a mean sample size of 42 and a median sample size of 35 for this design for the true DLT and response rates in Table 8. 10+10 accelerated titration or 15+15 accelerated titration designs yield even smaller mean sample sizes and result in less accuracy in dose selection, and hence were not considered further. Our results for the 20+20 accelerated titration design, which considers both drug toxicity and efficacy in selecting an optimal dose, are shown in Table 12.

Our simulation results in Table 12 for the 20+20 accelerated titration design show that for the true DLT and response rates in Table 8, dose level 4 is the dose that satisfies the Bayesian decision rules for safety and efficacy most frequently. Dose level 4 is chosen as the dose level that is acceptable for safety and efficacy in ~76% of the simulation runs in this case (also see Figure 9). In an actual trial, if more than one dose level satisfies these two Bayesian decision rules, other criteria such as the value of the utility function at each dose level, the percentage of patients at each dose level with a response but no DLT and the OR_i can be used to choose a single dose level, as mentioned earlier. In our example, dose level 4 has the maximum value of the utility function most frequently and has the maximum value for the average percentage of patients with a response but no DLT. The minimum value for the OR_i from the simulations is at dose level 1, not at dose level 4, but there are very few responders in

dose level 1 (see the first footnote to Table 12). Hence, based on these results of the 20+20 accelerated titration design, our final selection for optimal dose is dose level 4 for this example.

Results for the 20+20 accelerated titration design for various other scenarios of true toxicity and efficacy rates are shown in the Appendix.

In general, the results in Table 12 can be investigated for different values of the correlation coefficient r. However, r is restricted and can take on only certain values. For the example in Table 12, the highest positive correlation that can be used for the 10000 simulation runs to yield results with no errors due to the chosen value of r is 0.08, and those results do not differ substantially from the results shown in Table 12 for r=0. The highest value that r can take will differ for different combinations of true DLT and response rates at each dose (see the Appendix). For a logistic dose-toxicity curve and for a monotonically increasing dose-response curve, a higher value of r can be used (see the Appendix).

Comparison of the 20+20 Accelerated Titration Design to the Eff-Tox Design and to a 3+3 Design Followed by a Phase 2

We compare the performance of our proposed 20+20 accelerated titration design to that of the Eff-Tox design for various scenarios of true DLT and response rates that are provided as examples in the Eff-Tox program, as shown in Table 13. We do not compare those examples that are provided in the Eff-Tox program for which the

highest toxicity rate was very low (<=0.1) since the 20+20 part of our design is intended to target a DLT rate closer to 0.2. We use a sample size in the Eff-Tox design of 99 and cohort size of 9, which is the maximum cohort size allowed in the version of the Eff-Tox design software program we used. We also use 0.33 for the probability of the upper limit of toxicity and 0.5 for the probability of the lower limit of efficacy, identical to the values we use in the same decision rules in the 20+20 accelerated titration design at the end of each simulated trial. All the other input parameters used in the Eff-Tox design simulations are the same as those given in the Appendix as the Input parameters for Table 11. We also compare the performance of our proposed 20+20 accelerated titration design to that of the OBD Isotonic design by Zang et al., for the same scenarios of true DLT and response rates used in the comparison of the 20+20 accelerated titration design and the Eff-ox design. All the input parameters used in the OBD Isotonic design simulations are given in the Appendix.

For scenario 1 of Table 13, the Eff-Tox design does not select the dose with the highest trade-off value between the probabilities of efficacy and toxicity (i.e. dose level 3) most frequently as the optimal dose for efficacy and toxicity. The Eff-Tox design selects dose level 3, 41% of the time, a little less frequently than it selects dose level 4 (46% of the time), which has a slightly lower trade-off value. The OBD Isotonic design selects dose level 3, 38% of the time, less frequently than it selects dose level 4 (48% of the time). The 20+20 accelerated titration design selects dose level 3 and dose level 4 in 96% and 86% of the simulation runs respectively, as acceptable for

toxicity and efficacy. From our simulations of the 20+20 accelerated titration design, dose level 3 also has a) the maximum value of the utility function most frequently (39%, 53% and 59% of the time for c=0.1, c=0.5 and c=1 respectively), b) the maximum value for the average percentage of patients with a response but no DLT and c) the minimum value of the OR_i. Hence, based on these results of the 20+20 accelerated titration design, our final selection for this design is dose level 3.

For scenario 2, the Eff-Tox design selects dose level 1, the dose with the highest trade-off value, 74% of the time, the OBD Isotonic design selects dose level 1 53.1% of the time as the optimal dose. The 20+20 accelerated titration design selects dose level 1 as acceptable for safety and efficacy in ~94 % of the simulation runs. From our simulations of the 20+20 accelerated titration design, dose level 1 also has a) the maximum value of the utility function most frequently (53%, 72% and 85% of the time for c=0.1, c=0.5 and c=1 respectively), b) the maximum value for the average percentage of patients with a response but no DLT and c) the minimum value of the OR_i. Hence, based on these results of the 20+20 accelerated titration design, our final selection for this design is dose level 1.

For scenario 3, the Eff-Tox design does not select any dose as optimal for safety and efficacy 93% of the time, while the 20+20 accelerated titration design does not select any dose level as acceptable for safety and efficacy in 78% of the simulation runs. These results for the Eff-Tox design and the 20+20 accelerated titration design of

selecting no dose level as optimal or acceptable for safety and efficacy most of the time are reasonable, since the true DLT rate at the lowest dose level itself is quite high — at a value of 0.3, it is just below the upper limit of 0.33 considered in the Bayesian decision rule for safety. The OBD Isotonic design does not perform well in this scenario since it always assumes that at least the lowest dose should be safe, and selects dose level 1 and dose level 2 as the optimal dose 41% and 42% of the time respectively.

Although the comparisons between the 20+20 accelerated titration design and the Eff-Tox and OBD Isotonic designs are not exact in terms of cohort size and sample size, these examples demonstrate that the 20+20 accelerated titration design can select an optimal dose for efficacy and toxicity as robustly as the Eff-Tox design and the OBD Isotonic design but is easier to implement.

We also compare the results of the 20+20 accelerated titration design to those of a 3+3 Phase 1 design followed by a Phase 2 design (Table 14). The 3+3 design picks the right dose for safety i.e. the true MTD of dose level 4 only 60% of the time for the true DLT rates in Table 8, as seen in Table 9. Hence the probability of selecting the right dose for both toxicity and efficacy at the end of Phase 2 is no more than 60%. In contrast, as seen in Table 12, the 20+20 accelerated titration design picks dose level 4 as acceptable for safety and efficacy in ~76% of the simulation runs (76% is also the value for the percentage of simulations that select dose level 4 as the optimal dose

for toxicity and efficacy using the maximum value of the utility function for c=0.1) for the true toxicity and true response rates in Table 8. Since the 3+3 design has a small sample size, the accuracy of MTD selection is very low. Our simulations with this and other scenarios confirm that we need a larger sample size to select the right safe dose with high accuracy before proceeding to a Phase 2 study, so that an optimal dose for efficacy and toxicity can be selected with a high probability at the end of Phase 2.

Discussion and Conclusion

Most Phase 1 dose-finding oncology trials enroll a very small number of patients, and often fail to predict the MTD accurately due to the small sample size. We have shown, via examples of rule-based and model-based designs, that the accuracy of MTD or dose selection in these designs increases considerably with an increase in sample or cohort size. Thus, it is crucial to study the effect of sample size and cohort size on the accuracy of dose selection while designing an early phase oncology trial. With a larger number of patients, the efficacy of the drug can also be assessed in an early phase trial itself.

This has led us to propose a simple rule based design that enrolls a larger sample size than standard Phase 1 designs, that enables accurate dose selection with respect to toxicity and incorporates Bayesian decision rules for optimal dose selection for safety and efficacy at the end of the trial. In particular, we propose the “20+20 accelerated titration” design, a moderately sized, integrated Phase 1/2 trial design that assesses

both safety and efficacy. We note that such a design should not become too large; the drawbacks of large seamless Phase 1/2 trials with greater than a few hundred patients have been discussed by Mullard [111]. The 20+20 accelerated titration design is intended to quickly move up the dose levels through accelerated titration but to have large enough sample sizes for doses near the MTD to substantially increase the accuracy of MTD selection and to provide assessment of efficacy in treatment response. As mentioned, the 20+20 design targets a DLT rate of ~0.2, but our proposed 20+20 accelerated titration design has additional safety stopping rules which decrease the target DLT rate of the 20+20 part of the design. The stopping rules and their timing used in our proposed 20+20 accelerated titration design may be altered to create a modified 20+20 accelerated titration design. For example, one could implement stopping if there are >=3 DLTs, rather than >=4 DLTs, in the first 6 patients in a dose level, or one could apply stopping rules after patient 7, 14, 20 etc. instead of after patient 6, 14, 20 etc. However, this would further change the approximate DLT rate that the 20+20 part of the design targets, and simulations to study the operating characteristics of this modified design should be performed. In this context, we note that there is a trade-off for using aggressive stopping rules and stopping too early for toxicity – it decreases the probability of identifying the optimal dose. A safety committee that can stop the trial at its discretion at any point should be implemented in all these designs. In our simulations, we generate the true DLT rate at each dose level using a logistic dose-toxicity curve; we generate the true response rate at each dose level manually, since the efficacy of an anti-cancer drug may not

always increase with an increase in dose. Our simulations demonstrate that the 20+20 accelerated titration design can robustly pick a dose that is optimal for efficacy and safety.

We have also created a modified 20+20 accelerated titration design, where the switch from the 3+3 design to the 20+20 design occurs when either 1 DLT (grade 3, grade 4 or grade 5 toxicity) is observed or when 3 low grade toxicities (grade 1 or grade 2) are observed, similar to the switch from the 3+3 design to the CRM design in the paper by Iasonos [112]. The results for this modified design are very close to those from the 20+20 accelerated titration design in the settings that we have considered. Hence, we have provided results only for the 20+20 accelerated titration design.

The Eff-Tox design is an early phase design for oncology trials that considers both drug toxicity and efficacy. The Bayesian decision rules used in the Eff-Tox design to assign each new cohort of patients a dose imply that the dose that best optimizes efficacy and safety is selected each time. We employ these same decision rules, but only at the end of the trial in the 20+20 accelerated titration design. Hence, we have compared the performance of the proposed 20+20 accelerated titration design to that of the Eff-Tox design for some scenarios of true underlying DLT and response rates. We also compared the performance of the 20+20 accelerated titration design to that of the OBD Isotonic design proposed by Zang et al. for the same scenarios. Our comparisons show that the 20+20 accelerated titration design performs as well as or

better than the Eff-Tox design and the OBD Isotonic design for the scenarios considered. Our simple rule-based design can also be implemented more easily than the Eff-Tox design and the OBD Isotonic design.

We have also compared the results from the 20+20 accelerated titration design to those from a 3+3 Phase 1 design followed by a Phase 2 design. Our simulations confirm that we need a larger sample size to select the right safe dose with high accuracy before proceeding to a Phase 2 study so that a dose that is optimal for efficacy and toxicity can be picked with a high probability at the end of Phase 2. The importance of selecting the right safe dose and the optimal dose for toxicity and efficacy has been illustrated in several cases. As one example, it is illustrated by Markman [94] with the drug PLD, which is used to treat platinum resistant ovarian cancer. The dose of PLD explored in initial clinical trials as well as the dose approved by the FDA was 50 mg/m^2 (every 28 days). However clinical experience has shown that this dose leads to substantial adverse effects while a lower dose of 40 mg/m^2 is equally efficacious but leads to fewer adverse effects.

Table 7 Parameters Used to Determine the Coefficients of the Logistic Dose-Toxicity Curve

Parameter	Value
Starting dose	100 units (the remaining doses follow the modified Fibonacci series)
True DLT rate at starting dose	0.01
True MTD	Dose Level 4 = 501 units
True DLT rate at MTD	0.2

Table 8 True Underlying DLT and Response Rates

Dose	True probability of toxicity at each dose, generated from a logistic curve, whose coefficients are calculated using the parameters in Table 7 Log_e(DLT rate/(1-DLT rate)) =-5.39533+ 0.008002*dose	True probability of response at each dose is selected manually
100 units	0.01	0.01
200	0.02	0.05
334	0.06	0.15
501	0.2	0.45
701.4	0.55	0.2
932.86	0.89	0.05

Table 9 Accuracy of MTD Selection in Some Rule-Based Designs and for Some Cases for the CRM Design

Design	Target DLT Rate	% of Times the True MTD (Dose Level 4) is Selected	Median Sample Size, Maximum Sample Size
Standard 3+3 design	0.17–0.26	**60.0%**	**15, 33**
5+5 a design	0.2–0.25	**65.9%**	**30, 50**
CRM 5, 50 (cohort size of 5, sample size of 50)	0.15–0.25	**81.6%**	50 (maximum sample size, which is an input)
10+10 design	0.2–0.24	**74.0%**	**50, 80**
20+20 design	0.2–0.21	**90.1%**	**100, 140**
CRM 20,120 (cohort size 20, sample size 120)	0.15–0.25	**90.0%**	120 (maximum sample size, which is an input)

Table 10 Accuracy of MTD Selection in the CRM Design

Number of Patients on the Trial (Maximum Sample Size)	Cohort Size	% of Times the True MTD (Dose Level 4) is Selected
30	5	**69.2%**
40	5	**74.5%**
50	5	**81.6%**
60	5	**82.9%**
50	1	**81.4%**
50	2	**80.9%**
50	5	**81.6%**
120	20	**90.0%**
140	20	**92.8%**
160	20	**95.7%**
120	1	**95.3%**
120	2	**94.5%**
120	4	**94.5%**
120	5	**94.8%**
120	10	**94.1%**
120	20	**90.0%**

Table 11 Accuracy of Dose Selection in the Eff-Tox Design

Dose Level	1	2	3	4	5	None
True toxicity rate, true response rate	0.05, 0.2	0.1, 0.4	0.15, 0.6	0.2, 0.8	0.25, 0.95	
Utility (Trade-off Value)	**-0.68**	**-0.37**	**-0.04**	**0.28**	**0.51**	
% dose level selected Maximum Sample size=18 Cohort size=3	**2**	**4**	**31**	**34**	**29**	**1**
% dose level selected Maximum Sample size=99 Cohort size=3	**0**	**1**	**28**	**35**	**35**	**1**
% dose level selected Maximum Sample size=99 Cohort size=9	**0**	**1**	**26**	**23**	**49**	**0**

Table 12 Results for the 20+20 Accelerated Titration Design

	Dose Level 1	Dose Level 2	Dose Level 3	Dose Level 4	Dose Level 5	Dose Level 6
True toxicity, true efficacy rate	0.01, 0.01	0.02, 0.05	0.06, 0.15	0.2, 0.45	0.55, 0.2	0.89, 0.05
Odds of Toxicity to Efficacy from Simulations at Dose Level i (OR_i)*	**0.20**	**0.26**	**0.35**	**0.31**	**7.71**	**20.05**
Average Number of Patients	**3.5**	**4.5**	**7.3**	**14.0**	**12.2**	**0.28**
Average Number of DLTs	**0.04**	**0.1**	**0.45**	**2.8**	**6.7**	**0.25**
Average Number of Responses	**0.04**	**0.22**	**1.1**	**6.3**	**2.4**	**0.01**
Average Number of Patients with a Response but no DLT	**0.04**	**0.22**	**1.0**	**5.0**	**1.1**	**0.0**
Decision based on the Bayesian Posterior Probabilities (% of times out of 10000 simulations each dose satisfies the Bayesian decision criteria for both toxicity and efficacy)**	**2.86%**	**13.02%**	**28. 6%**	**76.08%**	**1.87%**	**0%**
% of times out of 10000 simulations each dose level is the one with the maximum value of the utility function						
when c =0.1***	**1.54%**	**4.46%**	**11.91%**	**76.15%**	**5.93%**	**0.01%**
when c = 0.5	**4.6%**	**6.66%**	**16.05%**	**71.63%**	**1.06%**	**0%**
when c=1	**7.34%**	**8.64%**	**20.79%**	**63.04%**	**0.19%**	**0%**

* Calculated using only those simulations runs with a non-zero denominator for OR_i. For the lower dose levels (levels 1 and 2), the denominator is zero in many simulation runs since the average number of responses is zero.

** No dose level is chosen as acceptable for toxicity and efficacy ~15% of the time. The addition of the percentages for dose selection based on the Bayesian decision rules can add up to more than 100 since more than one dose level can be chosen as acceptable for toxicity and efficacy in each simulation.

***c=1 gives equal weight to toxicity and efficacy while 0.1 gives a very small weight to toxicity and more weight to efficacy.

Mean sample size for this example is 41.75; median sample size is 35, minimum sample size is 12 and maximum sample size is 126.

Table 13 Comparing the Results of the Eff-Tox Design and the OBD Isotonic Design to those of the 20+20 Accelerated Titration Design for Various Scenarios of True Toxicity and Efficacy Rates

Scenario		Dose Level					
		1	2	3	4	5	None
1							
	True toxicity, true efficacy rate	.05, .10	.10, .30	.15, .60	.30, .62	.45, .65	-
	Trade-off Value Eff-Tox Design	**-0.88**	**-0.57**	**-0.04**	**-0.24**	**-0.41**	**-**
	% selected by Eff-Tox Design	**0**	**0**	**41**	**46***	**11**	**2**
	% selected by OBD Isotonic Design	**3.5**	**2.1**	**38.2**	**48.2**	**8**	**0**
	% of times each dose level is acceptable for efficacy and toxicity in the 20+20 Accelerated Titration Design (% of times each dose level is the one with the maximum value of the utility function when c =1)	**23.03** (4.88)***	**49.59 (15.64)**	**95.51 (58.81)**	**86.07 (18.85)**	**29.61 (2.13)**	**0.47**
2							
	True toxicity, true efficacy rate	.20, .60	.40, .62	.55, .65	.70, .70	.85, .75	-
	Trade-off Value Eff-Tox Design	**-0.12**	**-0.40**	**-0.57**	**-0.70**	**-0.83**	**-**
	% selected by Eff-Tox Design	**74**	**16**	**2**	**1**	**0**	**7**
	% selected by OBD Isotonic Design	**53.1**	**41.3**	**5.1**	**0.5**	**0**	**0**

	% of times each dose level is acceptable for efficacy and toxicity in the 20+20 Accelerated Titration Design (% of times each dose level is the one with the maximum value of the utility function when c =1)	**93.77** (84.64)*****	**60.27 (14.63)**	**5.59 (0.71)**	**0.02 (0.02)**	**0 (0)**	**3.52**
3							
	True toxicity, true efficacy rate	.30, .10	.40, .30	.55, .60	.60, .62	.65, .65	-
	Trade-off Value Eff-Tox Design	**-1.29**	**-1.04**	**-0.67**	**-0.71**	**-0.73**	**-**
	% selected by Eff-Tox Design	**0**	**3**	**4**	**0**	**0**	**93**
	% selected by OBD Isotonic Design	**40.7**	**41.6**	**15.3**	**2.2**	**0.2**	**0**
	% of times each dose level is acceptable for efficacy and toxicity in the 20+20 Accelerated Titration Design	**9.30****	**13.04**	**4.08**	**0.15**	**0**	**77.75**

*** The dose shown in green is the dose level selected by each design as the optimal dose for toxicity and efficacy.**

**** These numbers for percentage of times each dose level is acceptable for toxicity and efficacy can add up to more than 100% in the 20+20 accelerated titration design. For example, in scenario 1, dose levels 3 and 4 are selected in 96% and 86% of the simulation runs. This means that both dose levels 3 and 4 are selected in a large percentage of the 10000 simulations because both doses satisfy the Bayesian decision rules in those simulations.**

***** The percentages shown in brackets are the percentages that each dose is chosen as the optimal dose for the 20+20 accelerated titration design using the utility function, and these percentages add up to 100.**

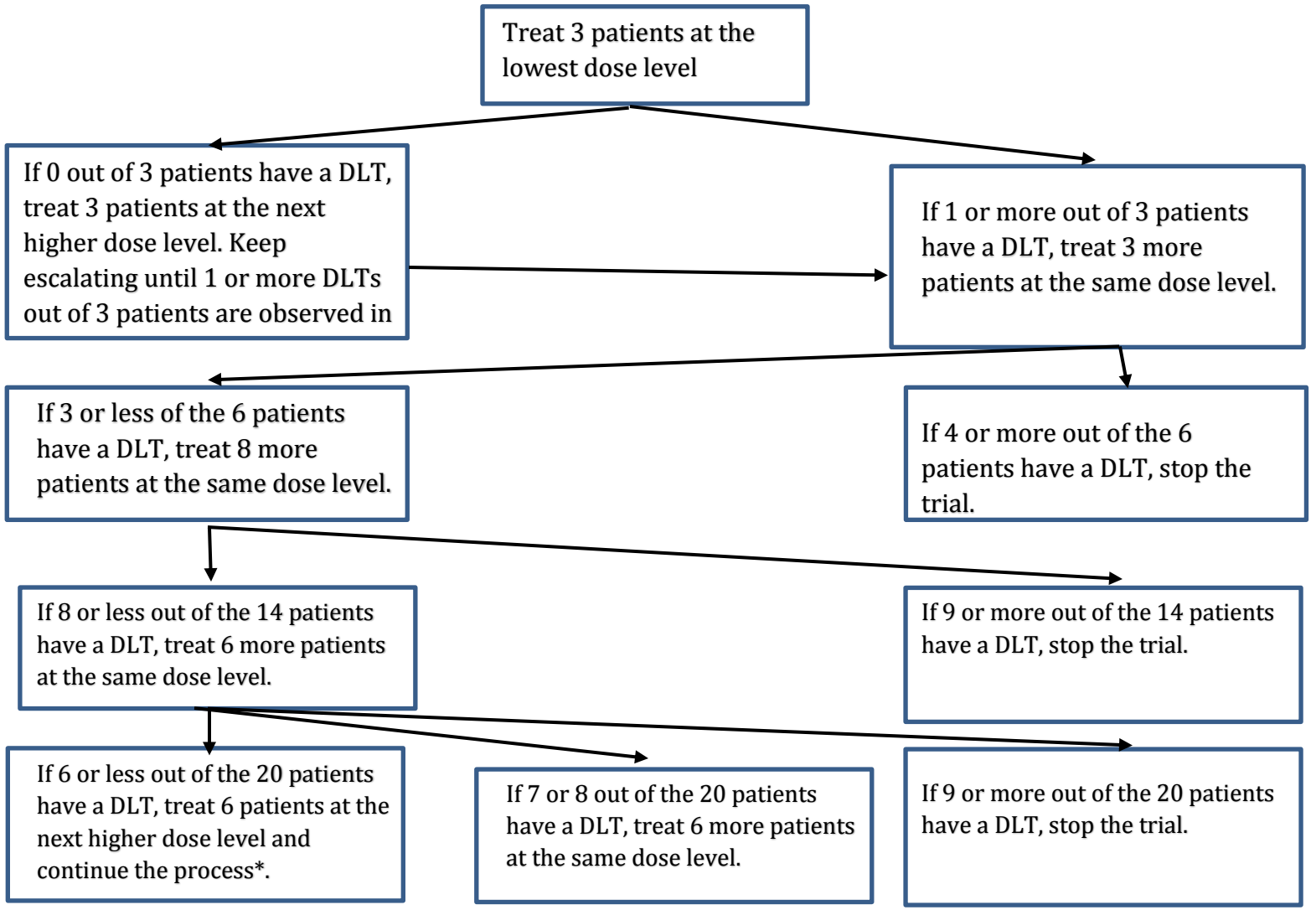
Table 14 Results for a 3+3 Phase 1 Design Followed by a Phase 2 Design

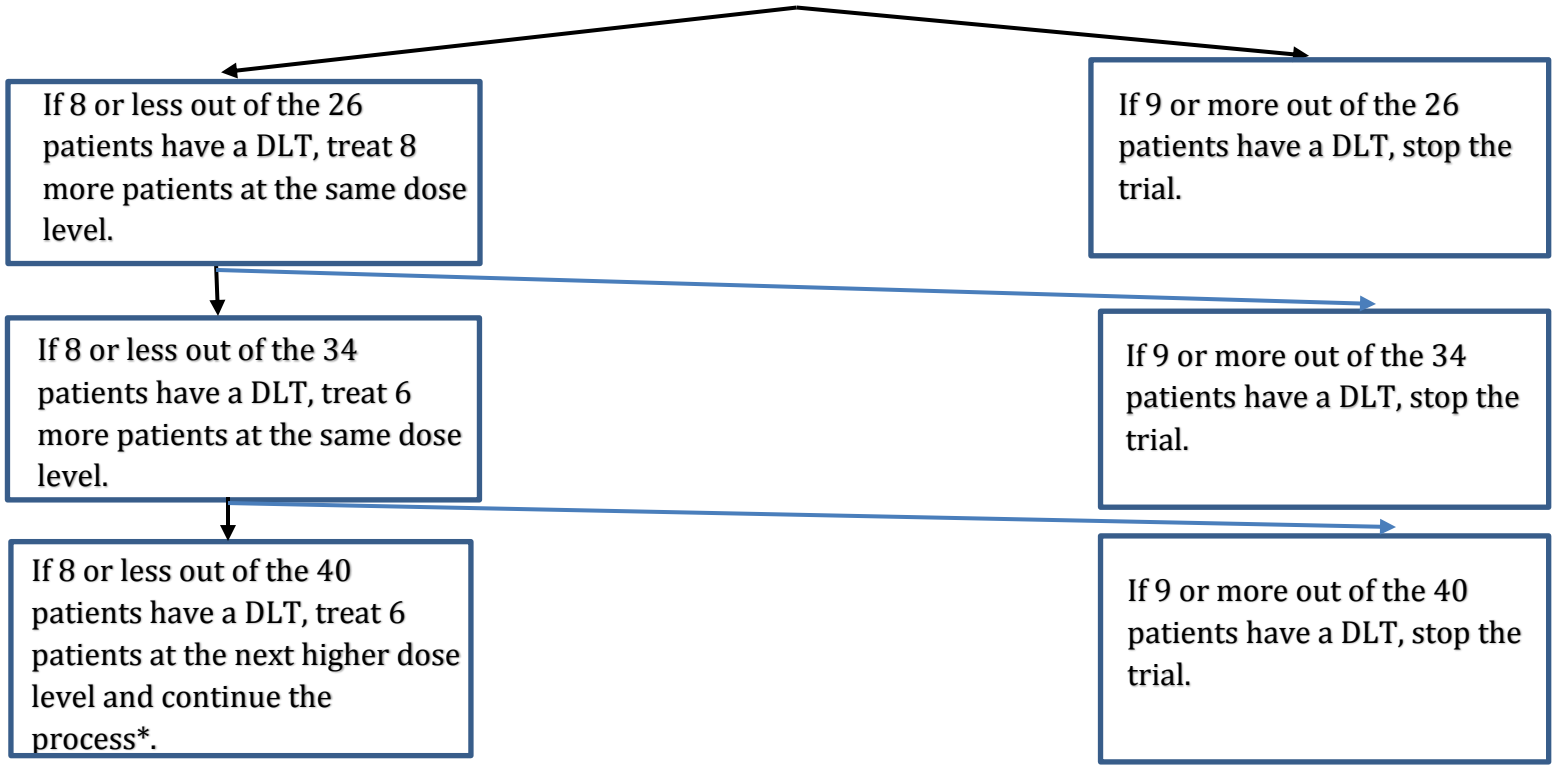
3+3 Phase 1 followed by Phase 2; Sample Size of 3+3 Design in Phase 1	**Hypothesis in Phase 2**	**Sample Size for Phase 2**	**Total Sample Size**	**Probability of Choosing the Right Dose for Toxicity and Efficacy**
Sample Size is ~15 (median) and ~33 (max)	H0=0.4, H1=0.6* Power=0.9	**52** (1-sided alpha =0.05) **64** (1-sided alpha = 0.025)	**~85 (max) (33 patients for 3+3 from Table 9 + 52 for Phase 2)** **~97 (max)**	**0.6*0.9=0.54 (0.6 is the probability of choosing the right dose for toxicity from Table 9) and 0.9 is the power in Phase 2 for the efficacy endpoint.**
Sample Size is ~15 (median) and ~33 (max)	H0=0.45, H1=0.65 Power=0.9	**52** (1-sided alpha =0.05) **63** (1-sided alpha = 0.025)	**~85 (max)** **~96 (max)**	**0.6*0.9=0.54**
Sample Size is ~15 (median) and ~33 (max)	H0=0.5, H1=0.7 Power=0.9	**50** (1-sided alpha =0.05) **62** (1-sided alpha = 0.025)	**~83 (max)** **~95 (max)**	**0.6*0.9=0.54**

***H0: response rate under the null hypothesis; H1: response rate under the alternative hypothesis**

For these calculations, we used the software EAST with the option "Discrete" for endpoint and "One Sample" for procedure. The option in EAST that says "Perform Exact Computations" was not used.

Figure 8 Schematic of 20+20 Accelerated Titration Design





* Note that we can escalate if we see 6 or less DLTs out of 20 patients or 8 or less DLTs out of 40 patients in a dose level, but we can also escalate if 0 DLTs and 0 responses are observed in the first 14 patients in a dose level since we cannot observe more than 6 DLTs in the last 6 patients and since the dose is not efficacious (otherwise we enroll the next 6 patients at the same dose level and continue the process).

Figure Legend: Schematic of the 20+20 accelerated titration design. At the end of the trial, the dose that is optimal for safety and efficacy is chosen using Bayesian decision rules and other criteria.

Figure 9 Percentage of Times Each Dose Level is Selected as Acceptable for Safety and Efficacy in the 20+20 Accelerated Titration Design Based on the DLT and Response Rates in Table 8

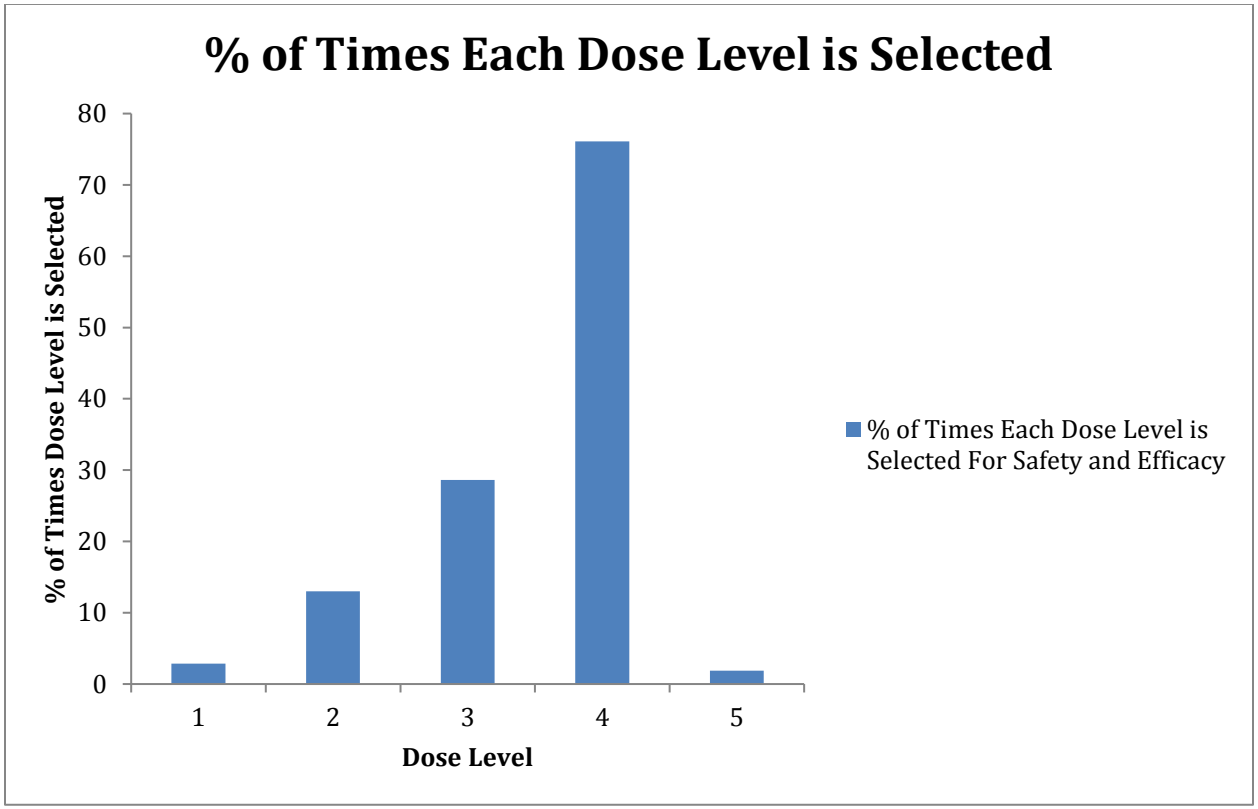


Figure Legend: Percentage of times out of 10000 simulations that each dose level is selected as acceptable for safety and efficacy in the 20+20 accelerated titration design based on the DLT and Response Rates in Table 8.

Appendix Chapter 3

1. Target DLT interval of the 20+20 design

We have constructed the stopping rules for the 20+20 design such that it targets a DLT rate Γ of ~0.2, similar to the 3+3 design, which targets an approximate DLT rate of 0.2 (range of 0.17–0.26) [80, 113]. Other stopping rules and cohort sizes can be proposed in order to target other DLT rates. The approximate DLT interval that any A+B design targets can be calculated using the following inequality from Ivanova [80].

$C_U/(A+B) < \Gamma < \Gamma_{A+B}$,

where Γ_{A+B} is the solution to the equation $Pr(Bin(A+B, \Gamma_{A+B})<=C_U)=0.5$.

A+B is the total number of patients a dose level can enroll and the trial is stopped if $>C_U$ DLTs are observed in A+B patients.

For our 20+20 design, A+B=40, C_U=8, and the inequality becomes:

$0.2<\Gamma<0.21$.

2. Correlation Coefficient r that Can Be Used (for Different Cases of True Efficacy and Toxicity Rates)

Dose Level	True toxicity rate	True efficacy rate
1	0.01	0.05
2	0.02	0.25
3	0.06	0.3
4	0.2	0.35
5	0.55	0.4
>=6	0.89	0.5

Maximum positive value of r in this case where both the true DLT and response rates are increasing monotonically with an increase in dose is ~0.25.

Dose Level	True toxicity rate	True efficacy rate
1	0.01	**0.4**
2	0.02	**0.35**
3	0.06	**0.3**
4	0.2	**0.25**
5	0.55	**0.15**
>=6	0.89	**0.05**

Maximum positive value of r in this case where the true DLT is increasing monotonically with an increase in dose but the true response rate is monotonically deceasing is 0.08.

3. Input Parameters Used in Simulations for the CRM, Eff-Tox and OBD Isotonic Designs

Input Parameters Used in the R Package CRM for the CRM Design

A CRM design with a target DLT rate of 0.2, starting dose level of 1 and a 1-parameter logistic dose-toxicity model with parameter "a" whose initial value is 1 and fixed parameter "b" whose value is 3 is considered. The prior for "a" is exp(-a). The prior DLT rate at each of the six dose levels is (0.15, 0.25, 0.3, 0.45, 0.51, 0.56) and the true DLT rate at each dose level is as given in Table 8.

Input parameters Used in the Eff-Tox Package for the Eff-Tox Design

Probability of Toxicity and Efficacy Limits for Dose Acceptability Rules

Parameter	Value
Prob(tox) upper limit (π_T*)	0.30000
Lower prob cutoff for prob of toxicity ($p_{T,L}$)	0.10000
Prob(eff) lower limit (π_E*)	0.50000
Lower prob cutoff for prob of efficacy ($p_{E,L}$)	0.10000

Trade-off Function Elicited Points (3 points to define the trade-off function contour)

	π_E	π_T
($\pi_{1,E}$*, 0)	0.50000	0.00000
(1, $\pi_{2,T}$*)	1.00000	0.65000
($\pi_{3,E}$, $\pi_{3,T}$)	0.70000	0.25000

Elicited Means (Prior Toxicity, Prior Efficacy)

Dose	Toxicity	Efficacy
1	0.0200	0.2000
2	0.0400	0.4000
3	0.0600	0.6000
4	0.0800	0.8000
5	0.1000	0.9000

Input parameters Used in the OBD Isotonic Design

Cohort size = 9

Number of cohorts = 11

phi = upper bound of toxicity rate = 0.33

ct = threshold for posterior probability of toxicity (any dose with toxicity probability larger than ct is excluded from the admissible set of doses) = 0.9

Number of simulations = 10000

The true DLT and response rates at each dose level are as given in Table 13.

The R code given at the following URL was used along with the input parameters given above:

http://odin.mdacc.tmc.edu/~yyuan/Software/TargetAgent/targetAgentDF.r

4. Results for the 20+20 Accelerated Titration Design for Various Scenarios of True Toxicity and Efficacy Rates

a) True Toxicity and Efficacy Rates are Monotonically Increasing with an Increase in Dose

	Dose Level 1	Dose Level 2	Dose Level 3	Dose Level 4	Dose Level 5	Dose Level 6
True toxicity, true efficacy rate	0.01, 0.01	0.02, 0.05	0.06, 0.15	0.2, 0.45	0.55, 0.5	0.89, 0.6
Odds of Toxicity to Efficacy from Simulations at Dose Level i (OR_i)*	**0.20**	**0.26**	**0.35**	**0.31**	**2.6**	**3.7**
Average Number of Patients	**3.5**	**4.5**	**7.3**	**14.0**	**12.2**	**0.28**
Average Number of DLTs	**0.04**	**0.1**	**0.45**	**2.8**	**6.7**	**0.25**
Average Number of Responses	**0.04**	**0.22**	**1.1**	**6.3**	**6.1**	**0.17**
Average Number of Patients with a Response but no DLT	**0.04**	**0.22**	**1.0**	**5.0**	**2.7**	**0.02**
Decision based on the Bayesian Posterior Probabilities (% of times out of 10000 simulations each dose level satisfies the Bayesian decision criteria for both toxicity and efficacy)**	**2.86%**	**13.02%**	**28. 6%**	**76.08%**	**15.32%**	**0%**
% of times out of 10000 simulations each dose level is the one with the maximum value of the utility function						
when c =0.1***	**0.31%**	**1.61%**	**6.43%**	**47.01%**	**42.91%**	**1.73%**
when c = 0.5	**2.13%**	**5.64%**	**14.04%**	**62.19%**	**15.63%**	**0.37%**
when c=1	**6.05%**	**8.48%**	**20.21%**	**61.13%**	**4.11%**	**0.02%**

* Calculated using only those simulations runs with a non-zero denominator for OR_i. For the lower dose levels (levels 1 and 2), the denominator is zero in many simulation runs since the average number of responses is zero.

** No dose level is chosen as acceptable for toxicity and efficacy ~13% of the time. The addition of the percentages for dose selection based on the Bayesian decision rules can add up to more than 100 since more than one dose level can be chosen as acceptable for toxicity and efficacy in each simulation.

***c=1 gives equal weight to toxicity and efficacy while 0.1 gives a very small weight to toxicity and more weight to efficacy.

Mean sample size for this example is 41.75; median sample size is35; minimum sample size is 12 and maximum sample size is 126.

b) True Toxicity Rates are Monotonically Increasing with an Increase in Dose and True Efficacy Rates Increase with an Increase in Dose but Plateau after Dose Level 3

	Dose Level 1	Dose Level 2	Dose Level 3	Dose Level 4	Dose Level 5	Dose Level 6
True toxicity, true efficacy rate	0.01, 0.01	0.02, 0.05	0.06, 0.15	0.2, 0.45	0.55, 0.45	0.89, 0.45
Odds of Toxicity to Efficacy from Simulations at Dose Level i (OR_i)*	**0.20**	**0.26**	**0.35**	**0.31**	**3.13**	**6.95**
Average Number of Patients	**3.5**	**4.5**	**7.3**	**14.0**	**12.2**	**0.28**
Average Number of DLTs	**0.04**	**0.1**	**0.45**	**2.8**	**6.7**	**0.25**
Average Number of Responses	**0.04**	**0.22**	**1.1**	**6.3**	**5.5**	**0.12**
Average Number of Patients with a Response but no DLT	**0.04**	**0.22**	**1.0**	**5.0**	**2.4**	**0.01**
Decision based on the Bayesian Posterior Probabilities (% of times out of 10000 simulations each dose level satisfies the Bayesian decision criteria for both toxicity and efficacy)**	**2.86%**	**13.02%**	**28.6%**	**76.08%**	**13.57%**	**0.0%**
% of times out of 10000 simulations each dose level is the one with the maximum value of the utility function when c =0.1*** when c = 0.5 when c=1	 **0.46%** **2.45%** **6.43%**	 **2.15%** **6.03%** **8.55%**	 **7.32%** **14.71%** **20.44%**	 **54.07%** **65.77%** **61.93%**	 **35.03%** **10.94%** **2.65%**	 **0.97%** **0.10%** **0%**

* Calculated using only those simulations runs with a non-zero denominator for OR_i. For the lower dose levels (levels 1 and 2), the denominator is zero in many simulation runs since the average number of responses is zero.

**No dose level is chosen as acceptable for toxicity and efficacy ~13% of the time. The addition of the percentages for dose selection based on the Bayesian decision rules can add up to more than 100 since more than one dose level can be chosen as acceptable for toxicity and efficacy in each simulation.

***c=1 gives equal weight to toxicity and efficacy while 0.1 gives a very small weight to toxicity and more weight to efficacy.

Mean sample size for this example is 41.75; median is 35; minimum sample size is 12 and maximum sample size is 126.

CHAPTER FOUR

Introduction

Several dose finding oncology designs have been developed over the past years that are improvements over the 3+3 design in terms of accuracy of MTD selection as well as other operating characteristics such as the percentage of patients under-dosed. Here, we focus on two of these relatively recent designs: the TEQR and the mTPI designs. The TEQR design is a simple dose finding design where the dose finding decisions are based on whether the empirical dose limiting toxicity (DLT) rate lies within the target DLT interval or in the interval below or above it [73]. The mTPI design [72] uses a similar concept for dosing decisions but provides a Bayesian counterpart to the frequentist TEQR design.

Phase I trials are generally very small and the accuracy of MTD selection is low with such a small sample size. Hence, we first compare the frequentist TEQR dose-finding design and the Bayesian mTPI dose-finding design for accuracy of MTD selection with larger sample sizes and identical stopping rules. We then extend the TEQR and mTPI designs with a moderately large sample size to choose an optimal dose for both safety and efficacy by considering correlated Bernoulli distributions for the true underlying toxicity and efficacy rates; we have previously proposed a simple rule-based design that incorporates toxicity and efficacy in dose selection, and we now continue this work and similarly extend the TEQR and mTPI designs to also include efficacy. In our

simulations, we assume that the true DLT rates increase monotonically with an increase in dose but we do not assume that the true response rates increase monotonically with an increase in dose; we allow multiple types of curves for dose-response (monotonically increasing, plateau, or umbrella-shaped). In this context, we apply a recent update to the isotonic regression method, called nearly-isotonic regression, and investigate if it can be used to select a dose that is optimal for both safety and efficacy. We also apply isotonic regression on the difference in observed response rates between adjacent dose levels and investigate if it can be used to improve the accuracy of dose selection for certain dose-response curves. Thus, we propose a simple way of extending the mTPI (and TEQR) design to include efficacy, using the technique of isotonic regression. This is in contrast to the more complex technique of using a statistic called the joint unit probability mass (JUPM) for toxicity and efficacy, as proposed by Li et al., to extend the mTPI design to include efficacy [114]. We finally compare the accuracy of dose selection of the extended TEQR and mTPI designs to that of the Eff-Tox design and the Optimal Biological Dose (OBD) Isotonic design.

Methods

TEQR and MTPI Designs

The TEQR design is a frequentist design and is based on the empirical DLT rate [73]. The toxicity probability scale is divided into three intervals, namely $(0, p_T-\varepsilon_1)$, $[p_T-\varepsilon_1,$

$p_T+\varepsilon_2$] and ($p_T+\varepsilon_2$,1); p_T is the target probability of DLT and ε_1 and ε_2 are used to define the interval for the target DLT rate. The rules for escalating, staying at the same dose or de-escalating depend on which of these intervals contains the empirical DLT rate for that dose level – if the empirical DLT rate lies between 0 and $p_T-\varepsilon_1$, the next cohort of patients will be treated at the next higher dose; if it lies in the interval [$p_T-\varepsilon_1$, $p_T+\varepsilon_2$], the next cohort of patients will be treated at the same dose; if it lies above $p_T+\varepsilon_2$, the next cohort of patients will be treated at a lower dose. The trial stops if dose level 1 is too toxic or when a dose level achieves the selected MTD sample size.

The mTPI design is a Bayesian counterpart of the TEQR design and uses the unit probability mass (UPM) statistic, defined as the ratio of the probability mass of the interval and the length of the interval [72], for the dose finding decisions. The toxicity probability scale is again divided into three intervals, namely (0, $p_T-\varepsilon_1$), [$p_T-\varepsilon_1$, $p_T+\varepsilon_2$] and ($p_T+\varepsilon_2$, 1), and these three intervals correspond to under-dosing, correct dosing and over-dosing respectively. The rules for escalating, staying at the same dose or de-escalating depend on which of these intervals has the highest UPM for that dose level, based on a beta-binomial posterior distribution formed from the likelihood of the observed DLT data and a beta(1,1) prior. For example, the next cohort of patients will be treated at the same dose if the UPM is the largest for the correct dosing interval. The trial stops if dose level 1 is too toxic or if the maximum sample size is reached or exceeded. In both the mTPI and TEQR design, we stay at the current dose if the current dose is safe but the DLT data indicate that the next higher dose is too toxic.

Isotonic Regression

When the true underlying DLT rate (or response rate) increases with an increase in dose, the observed DLT (or response) rate is also expected to be a monotonically non-decreasing function of dose. However, this may not always be what is observed due to the small sample size in each dose level in dose-finding oncology trials. Isotonic regression is a weighted regression and a smoothing procedure that has been used to ensure that the estimated DLT (or response) rate is a monotonically non-decreasing function of dose. This then enables us to determine the highest dose level that is acceptable for safety and the lowest dose level that is acceptable for efficacy.

In a trial using a standard TEQR or mTPI design, the dose chosen for safety is the highest dose level with a DLT rate that is closest to (and below) the pre-specified DLT rate (say 0.33) after applying isotonic regression at the end of the trial. In our simulated trials of the TEQR or mTPI design that has been extended to evaluate both safety and efficacy, isotonic regression is also applied independently to the observed response rates at the end of each trial, when the true underlying response rates are thought to be monotonically increasing or monotonically non-decreasing with an increase in dose. Since the estimated response rates will be monotonically non-decreasing with an increase in dose after applying isotonic regression, we choose as the optimal dose for safety and efficacy the highest dose level where the DLT rate is less than or equal to 0.33 after isotonic regression, only if the smoothed response rate at that dose level crosses the efficacy threshold (say response rate of 0.4). For

example, if dose level 4 is chosen after isotonic regression as the **highest** dose level with a DLT rate <=0.33 and dose level 3 or lower is chosen after isotonic regression as the **lowest** dose level with a response rate >=0.4, then dose level 4 is the optimal dose for safety and efficacy since the response rate at dose level 4 will be >=0.4 in this case. However, if dose level 3 is chosen for safety after isotonic regression and dose level 4 is chosen for efficacy after isotonic regression, then there is no dose level that is optimal for safety and efficacy because the efficacy threshold of a response rate of 0.4 is not crossed at dose level 3, but only at dose level 4. Of course, if dose level 3 is chosen for both safety and efficacy after isotonic regression, then dose level 3 is the optimal dose for safety and efficacy.

Nearly-isotonic Regression

When the true response rates are not considered to be strictly monotonically increasing with an increase in dose, we can use the method of "nearly-isotonic" regression developed by Holger et al. of approximating a sequence of data points with a nearly-monotone function [115]. Their method is a modified version of the pool adjacent violators algorithm (PAVA), and the formulation of the method includes a parameter λ that controls the amount of smoothing of the data. In mathematical terms, the equation that is considered is the following convex optimization problem for each non-negative λ:

$$\hat{\beta}_\lambda = \underbrace{argmin}_{\beta \in R^n} \frac{1}{2} \sum_{i=1}^{n} (y_i - \beta_i)^2 + \lambda \sum_{i=1}^{n-1} (\beta_i - \beta_{i+1})_+$$

where $\hat{\beta}_\lambda$ (= (β_1, β_2,…, β_n)) is the nearly monotone approximation to the sequence of n points y_1, y_2,…..y_n, and x+ indicates the positive part, x+ = x · 1(x > 0). $\hat{\beta}_\lambda$ will vary with λ. When λ=0, then there is no smoothing at all and the estimates are fit to just the observed data points as is; as λ gets larger and larger, the regression fit tends towards the standard isotonic regression fit.

Finding the Peak of an Umbrella-shaped Dose-Response Curve using Isotonic Regression

When there is a peak in the dose-response curve (umbrella-shaped dose-response curve), we apply isotonic regression to the **difference** in observed response rates between adjacent dose levels obtained at the end of each simulated trial. These differences function like a derivative. For a convex shaped curve for example, the derivative is 0 at the peak, and the sign of the derivative changes from before the peak to after the peak. This is the same concept we use to determine the peak of an umbrella shaped dose-response curve – we apply isotonic regression at the end of each simulated trial to the differences in observed response rates between adjacent dose levels, and observe where these differences switch from a negative to a positive sign, to determine the peak of the curve.

Comparisons of Results for Accuracy of Optimal Dose Selection

We compare the results for accuracy of optimal dose selection of the extended TEQR and mTPI designs to those of the Eff-Tox design and the Optimal Biological Dose (OBD) Isotonic design for various scenarios of true toxicity and efficacy rates. The Eff-Tox design [67, 102] and the OBD Isotonic design [101] are described in Chapter 3.

Simulations

We generate two correlated binary random variables X1~ Bernoulli(p) and X2~ Bernoulli(q) for toxicity and efficacy respectively as follows. We first generate X1 with success probability p. If X1=1, we generate X2 with probability q1 and if X1=0, we generate X2 with probability q2. q1 and q2 are defined as follows:

q1 = q + (r/p) * sqrt(p*(1-p)*q*(1-q))

q2 = (q - q1*p)/(1-p)

r is the correlation coefficient between X1 and X2 and is restricted to lie between $\max(-(\frac{(pq)}{(1-p)(1-q)})^{\frac{1}{2}}, -(\frac{(1-p)(1-q)}{(pq)})^{\frac{1}{2}})$ and $\min((\frac{(p\,(1-q))}{(1-p)(q)})^{\frac{1}{2}}, (\frac{((1-p)q)}{(p)(1-q)})^{\frac{1}{2}})$ [110].

In the equations for q1 and q2, it can be seen that if the correlation coefficient r=0 then q1=q2=q.

We generate the true DLT rate at each dose level (p_i) using a logistic dose toxicity curve, whose two coefficients are calculated using the following parameters: true DLT rate at starting dose (dose level 1, 100 units) of 0.01 and true DLT rate of 0.2 at the

true MTD (dose level 4, 501 units). We generate the true response rate at each dose level (q_i) manually to follow either a monotonic increase, a plateau, or an umbrella-shape (Table 15).

We have created our own SAS code to simulate both the extended TEQR and mTPI designs. To obtain the statistical operating characteristics for each scenario in the results section, we perform 1000 simulations each time. The rules for escalation, de-escalation or remaining at the same dose for each simulated trial are based on the number of observed DLTs. Two different stopping rules are considered in our simulations, namely the usual stopping rules for the TEQR design and the mTPI design respectively; the simulated trial stops a) when the desired MTD sample size is reached or when dose level 1 is too toxic or b) when the desired total sample size is reached or when dose level 1 is too toxic. In our simulations of these designs, we have also considered the underlying true response rate at each dose level and the resultant response of each patient. Although the dose escalation/staying/de-escalation decisions are based only on the number of observed DLTs, the response of each patient is also observed and noted. At the end of each simulated trial, we choose a dose that is optimal for both safety and efficacy based on the observed DLT and response rates at each dose level.

The input parameters used in our SAS code for the TEQR and mTPI designs are provided in Appendix Table 1. For simplicity, the coefficient of correlation **r is set to**

0 for the simulation results given in the main text. The results can be investigated for correlation coefficients other than zero (Appendix Section 2) within the range of values that the correlation coefficient can assume.

Results

Safety Only: Monotonically Increasing True DLT Rates

Only the monotonically increasing true DLT rates with an increase in dose shown in Table 15 are used in the simulations and no efficacy is involved in the results in Tables 16 and 17; isotonic regression is applied to the observed DLT rates at the end of each simulated trial to determine the MTD.

We use the same stopping rules for the mTPI and TEQR designs and compare them for accuracy of MTD Selection.

a) We use the usual stopping rules of the TEQR design, namely stop the trial when the desired MTD sample size is reached or when dose level 1 is too toxic, for both the TEQR and mTPI designs and compare their performance for the accuracy of MTD selection; dose level 4 is the true MTD in our example.
b) We use the usual stopping rules of the mTPI design, namely stop the trial when the total desired sample size is reached or when dose level 1 is too toxic, for both the TEQR and mTPI designs and compare their performance for the accuracy of MTD selection; dose level 4 is the true MTD in our example.

In general, we find that when identical stopping rules are used for both the designs,

the Bayesian mTPI design is more accurate than the frequentist TEQR design in selecting the true MTD for the scenarios explored, with the same or similar number of subjects (Table 16 and Table 17). Also, given the same cohort size, the accuracy of MTD selection generally increases when the total sample size is increased. Thus, we use a moderate sample size of 50 subjects to evaluate efficacy and safety.

Safety and Efficacy: Monotonically Increasing True DLT and Response Rates

The monotonically increasing true DLT and response rates with an increase in dose shown in Table 15 are used in the simulations and isotonic regression is applied independently to the observed DLT rates and to the observed response rates at the end of each simulated trial.

a) We use the usual stopping rules of the TEQR design, namely stop the trial when the desired MTD sample size is reached or when dose level 1 is too toxic, for both the extended TEQR and mTPI designs and compare their performance for dose selection. The results in Table 18 are based on a MTD sample size of 50 and a cohort size of 5.

For this example, both the extended TEQR and mTPI designs select dose level 4 as the optimal dose for safety and efficacy with the highest probability (Table 18). The extended mTPI design selects dose level 4 as the optimal dose with a higher probability than the extended TEQR design does.

For the extended TEQR design example in Table 18, we also apply isotonic regression to the **difference** in the observed response rates between adjacent dose levels at the

end of each simulated trial, to investigate if this technique will help determine the dose for efficacy (Appendix Section 3).

b) We use the stopping rules of the mTPI design, namely stop the trial when the total desired sample size is reached or when dose level 1 is too toxic, for both the extended TEQR and mTPI designs and compare their performance for dose selection. The results in Table 19 are based on a total sample size of 50 and a cohort size of 5.

For this example, both the extended TEQR and mTPI designs select dose level 4 as the optimal dose for safety and efficacy with the highest probability (Table 19). The extended mTPI design selects dose level 4 as the optimal dose with a higher probability than the extended TEQR design does.

Safety and Efficacy: Monotonically Increasing True DLT Rates and Plateauing Response Rates

The monotonically increasing true DLT and plateauing response rates with an increase in dose shown in Table 15 are used in the simulations and isotonic regression is applied independently to the observed DLT rates and to the observed response rates at the end of each simulated trial.

a) We use the usual stopping rules of the TEQR design, namely stop the trial when the desired MTD sample size is reached or when dose level 1 is too toxic, for both the extended TEQR and mTPI designs and compare their performance for dose

selection. The results in Table 20 are based on a MTD sample size of 50 and a cohort size of 5.

For this example, both the extended TEQR and mTPI designs select dose level 4 as the optimal dose for safety and efficacy with the highest probability (Table 20). The extended mTPI design selects dose level 4 as the optimal dose with a higher probability than the extended TEQR design does.

For the extended TEQR design example in Table 20, we also apply isotonic regression to the **difference** in the observed response rates between adjacent dose levels at the end of each simulated trial, to investigate if this technique will help determine the dose for efficacy (Appendix Section 4).

b) We use the stopping rules of the mTPI design, namely stop the trial when the total desired sample size is reached or when dose level 1 is too toxic, for both the extended TEQR and mTPI designs and compare their performance for dose selection. The results in Table 21 are based on a total sample size of 50 and a cohort size of 5.

For this example, both the extended TEQR and mTPI designs select dose level 4 as the optimal dose for safety and efficacy with the highest probability (Table 21). The extended mTPI design selects dose level 4 as the optimal dose with a higher probability than the extended TEQR design does.

Safety and Efficacy: Monotonically Increasing True DLT Rates and Umbrella-Shaped Dose-Response Curve

The monotonically increasing true DLT rates with an increase in dose and the umbrella-shaped true response rates shown in Table 15, where the response rate peaks at dose level 4, are used in the simulations; isotonic regression is applied to the observed DLT rates and nearly-isotonic regression is applied to the observed response rates at the end of each simulated trial.

a) We use the usual stopping rules of the TEQR design, namely stop the trial when the desired MTD sample size is reached or when dose level 1 is too toxic, for both the extended TEQR and mTPI designs and compare their performance for dose selection. The results in Table 22 are based on a MTD sample size of 50 and a cohort size of 5.

In general, we will not know the exact shape of the underlying dose-response curve at the beginning of the trial. However, when the true dose-response curve is thought to be umbrella-shaped, the results in Table 22 are obtained when the technique of nearly-isotonic regression is applied to the observed response rates. When the true underlying response rates possess a peak at a certain dose, the dose level that is chosen for efficacy depends on the amount of smoothing performed on the observed response rates – this smoothing is controlled by the parameter λ. In this scenario, applying nearly-isotonic regression or isotonic regression directly to the response rates does not help identify the peak in the dose-response curve. When the true dose-

response curve is thought to possess a clear peak, we suggest applying isotonic regression to the **difference** in observed response rates between adjacent dose levels to reveal or identify this peak dose level for efficacy, as described in Appendix Section 5.

For the extended TEQR design example in Table 22, we also apply isotonic regression to the **difference** in the observed response rates between adjacent dose levels at the end of each simulated trial, to help determine the dose at which the response rate peaks (Appendix Section 5).

b) We use the stopping rules of the mTPI design, namely stop the trial when the total desired sample size is reached or when dose level 1 is too toxic, for both the extended TEQR and mTPI designs and compare their performance for dose selection. The results in Table 23 are based on a total sample size of 50 and a cohort size of 5.

In this case, where the true underlying response rates possess a peak at a certain dose, the dose level that is chosen for efficacy depends on the amount of smoothing done on the observed response rates – this smoothing is controlled by the parameter λ. In this scenario, applying nearly-isotonic regression or isotonic regression directly to the response rates does not help identify the peak in the dose-response curve. When the true dose-response curve is thought to possess a clear peak, we suggest applying isotonic regression to the **difference** in observed response rates between adjacent

dose levels to reveal or identify this peak dose level for efficacy, as described in Appendix Section 5.

Comparison of the Accuracy of Optimal Dose Selection for Various Designs

We compare the accuracy of optimal dose selection of our extended TEQR and mTPI designs to that of the Eff-Tox design and that of the OBD Isotonic design by Zang et al., for some scenarios of true DLT and response rates. All the input parameters used in the Eff-Tox design and OBD Isotonic design simulations are given in Appendix Section 6.

We find that among the designs considered, the extended mTPI design selects the optimal dose more accurately than the other designs for these scenarios. The extended TEQR design performs as well as or slightly worse than the Eff-Tox design in terms of accuracy of optimal dose selection in most scenarios. The OBD Isotonic design performs well for the case of the umbrella-shaped dose response curve, while the Eff-Tox design frequently does not select any dose as optimal for such a dose-response curve (Table 24).

Discussion and Conclusion

We have first compared the frequentist TEQR design with the Bayesian mTPI design for accuracy of MTD selection, when using the same stopping rules for both designs.

In all the scenarios considered, the Bayesian mTPI design is more accurate in selecting the true MTD than the frequentist TEQR design, when identical stopping rules and the same or similar sample sizes are used for both the designs. Also, given the same cohort size, the accuracy of MTD selection generally increases when the total sample size is increased.

We then extended the TEQR and mTPI designs to also consider efficacy, in addition to safety, in a moderately sized trial. In our extended TEQR or mTPI trial designs, isotonic regression is always applied to the observed DLT rates at the end of the trial, since the true DLT rate is always assumed to increase with an increase in dose. The technique that is most appropriate to apply to the observed response rates depends on the drug's properties (Figure 10). For this, clinical knowledge or judgement about the true underlying response rates of the study drug is required to have a good initial guess at the true dose-response curve.

When the true underlying response rates are thought to increase monotonically with an increase in dose or to first increase monotonically and then plateau after a certain dose level, isotonic regression can also be applied to the observed response rates at the end of the TEQR or mTPI trial. The optimal dose level for safety and efficacy is chosen to be the highest dose level for which the DLT rate after applying isotonic regression is below or at the target toxicity rate, only if the threshold for response rate is crossed at that dose. If the threshold for response rate is not reached at the

highest dose level at which the smoothed DLT rate is below or at the target toxicity rate, then no dose level is chosen as optimal for safety and efficacy.

When the underlying true response rates are thought to possess a clear peak (umbrella shaped dose-response curve), isotonic regression on the **difference** in observed response rates between adjacent dose levels, along with the sign of these differences, can be used to reveal or identify this peak dose level for efficacy. This information of the peak dose level for efficacy can then be used in conjunction with the dose level picked for safety, to select an optimal dose for safety and efficacy. For example, if the peak dose level identified for efficacy is equal to or lower than the dose level selected for safety, then the peak dose level identified for efficacy is chosen as the optimal dose for safety and efficacy, assuming that the peak is above the specified efficacy threshold – if not, no dose level is chosen as the optimal dose. If the peak dose level identified for efficacy is higher than the dose level selected for safety, then the lower dose level selected for safety can be chosen as the optimal dose, only if the response rate at that dose is greater than or equal to the efficacy threshold – if not, no dose is chosen as the optimal dose. In brief, we cannot select a dose that exceeds the target toxicity, but if the maximum or peak efficacy of the drug is reached at a lower dose, we can use that as the optimal dose assuming the efficacy threshold is crossed at that dose (Figure 11).

Note that when we use isotonic regression on the difference in response rates between adjacent dose levels in the case when there is no peak in the response rates (monotonically increasing or plateauing response rates), we find that no dose level is selected as the peak frequently. In this case, we can apply isotonic regression on the response rates themselves.

We compared the extended TEQR and mTPI design to the Eff-Tox and OBD Isotonic design for accuracy of optimal dose selection for some scenarios of true efficacy and toxicity rates. For these scenarios and for the designs considered, we found that the extended mTPI design selects the optimal dose more accurately than the other designs. The extended TEQR design performs as well as or slightly worse than the Eff-Tox design in terms of the accuracy of optimal dose selection for most of the scenarios. The OBD Isotonic design performs well for an umbrella-shaped dose response curve, while the Eff-Tox design frequently does not select any dose as optimal for such a dose-response curve.

In summary, we have continued to propose designs that incorporate toxicity and efficacy in dose selection. In this context, we applied a recent update to the isotonic regression method, called nearly- isotonic regression, and investigated if it could be used to select a dose that is optimal for both safety and efficacy. We did not find that it was necessary to use this update to isotonic regression but found that isotonic regression itself applied on the difference in observed response rates between

adjacent dose levels could be used to identify the peak of a dose-response curve with a clear maximum, such as a convex umbrella-shaped dose-response curve. For other dose-response curves, such as monotonically increasing or plateau, applying isotonic regression to both the observed DLT and response rates independently can be used to determine the optimal dose for toxicity and efficacy.

Table 15 Monotonically Increasing True DLT Rates with an Increase in Dose and Different Dose-Response Curves (the true probability of response at each dose is selected manually)

Dose	True probability of toxicity at each dose, generated from a logistic curve, whose coefficients are calculated assuming the true DLT rate at 100 units to be 0.01 and at 501 units to be 0.2. The dose levels follow the modified Fibonacci series. Log_e(DLT rate/(1-DLT rate)) =-5.39533+ 0.008002*dose	**Monotonically Increasing True Response Rates**	**Plateauing Response Rates with an Increase in Dose**	**Umbrella-Shaped Dose-Response Curve**
100 units	0.01	0.1	0.1	0.1
200	0.02	0.3	0.3	0.3
334	0.06	0.4	0.4	0.4
501	0.2	0.45	0.45	0.45
701.4	0.55	0.55	0.45	0.2
932.86	0.89	0.6	0.45	0.05

Table 16 Results for Accuracy of MTD Selection: Stopping Rules of the TEQR Design

Cohort Size	Sample Size at MTD	**mTPI Accuracy of MTD Selection (% of Times out of 1000 Simulations Dose Level 4 is Selected)**	**Median Sample Size**	**Maximum Sample Size**	**TEQR Accuracy of MTD Selection (% of Times out of 1000 Simulations Dose Level 4 is Selected)**	**Median Sample Size**	**Maximum Sample Sze**
3	30	**75.7%**	**39**	**75**	**73.1%**	**45**	**84**
4	40	**83.4%**	**56**	**100**	**69.5%**	**52**	**92**
5	50	**85.0%**	**65**	**115**	**63.7%**	**65**	**115**
6	60	**82.2%**	**78**	**132**	**82.5%**	**84**	**132**
10	100	**93.4%**	**130**	**200**	**83.7%**	**130**	**210**
Sample Size at MTD	Cohort Size	**mTPI Accuracy of MTD Selection (% of Times out of 1000 Simulations Dose Level 4 is Selected)**	**Median Sample Size**	**Maximum Sample Size**	**TEQR Accuracy of MTD Selection (% of Times out of 1000 Simulations Dose Level 4 is Selected)**	**Median Sample Size**	**Maximum Sample Sze**
60	2	**79.2%**	**70**	**132**	**43.4%**	**66**	**128**
60	3	**75.1%**	**69**	**147**	**75.9%**	**75**	**135**
60	4	**83.9%**	**76**	**140**	**66.6%**	**72**	**128**
60	5	**87.3%**	**75**	**135**	**62.1%**	**75**	**130**
60	6	**82.2%**	**78**	**132**	**82.5%**	**84**	**132**
60	10	**91.3%**	**90**	**130**	**81.6%**	**90**	**130**

Table 17 Accuracy of MTD Selection: Stopping Rules of the mTPI Design

Cohort Size	Total Sample Size	**mTPI Accuracy of MTD Selection (% of Times out of 1000 Simulations Dose Level 4 is Selected)**	**TEQR Accuracy of MTD Selection (% of Times out of 1000 Simulations Dose Level 4 is Selected)**
3	30	**71.5%**	**66.6%**
4	40	**80.3%**	**68.7%**
5	50	**86.2%**	**64.5%**
6	60	**82.7%**	**81.3%**
10	100	**91.5%**	**82.8%**

Total Sample Size	Cohort Size	**mTPI Accuracy of MTD Selection (% of Times out of 1000 Simulations Dose Level 4 is Selected)**	**TEQR Accuracy of MTD Selection (% of Times out of 1000 Simulations Dose Level 4 is Selected)**
60	2	**78.1%**	**44.1%**
60	3	**76.0%**	**74.7%**
60	4	**85.1%**	**67.2%**
60	5	**86.4%**	**66.6%**
60	6	**82.7%**	**81.3%**
60	10	**90.4%**	**79.7%**

Table 18 Results for Accuracy of Dose Selection in the Extended TEQR and mTPI Designs for Monotonically Increasing True DLT and Response Rates: Stopping Rules of the TEQR Design

TEQR Design

Dose Level	% of Times Dose Level is Chosen for Toxicity	% of Times Dose Level is Chosen for Efficacy	Probability of Dose Level being Selected as Optimal for Toxicity and Efficacy
No Dose Level is chosen	0%	9.2%	0.28
1	0.8%	0%	0
2	3.2%	27.9%	0.01
3	31.8%	28.1%	0.18
4	63.7%	26.6%	0.53
5	0.5%	8.1%	0
6	0%	0.1%	0

mTPI Design

Dose Level	% of Times Dose Level is Chosen for Toxicity	% of Times Dose Level is Chosen for Efficacy	Probability of Dose Level being Selected as Optimal for Toxicity and Efficacy
No Dose Level is chosen	0%	9.9%	0.21
1	0%	0%	0
2	0.2%	26.5%	0
3	14.7%	29.1%	0.08
4	85.0%	27.3%	0.70
5	0.1%	7.1%	0
6	0%	0.1%	0

Table 19 Results for Accuracy of Dose Selection in the Extended TEQR and mTPI Designs for Monotonically Increasing True DLT and Response Rates: Stopping Rules of the mTPI Design

TEQR Design

Dose Level	% of Times Dose Level is Chosen for Toxicity	% of Times Dose Level is Chosen for Efficacy	Probability of Dose Level being Selected as Optimal for Toxicity and Efficacy
No Dose Level is chosen	0%	9.4%	0.28
1	0.6%	0%	0
2	2.6%	25.4%	0.01
3	31.3%	32.3%	0.18
4	64.5%	24.0%	0.53
5	1.0%	8.9%	0.01

mTPI Design

Dose Level	% of Times Dose Level is Chosen for Toxicity	% of Times Dose Level is Chosen for Efficacy	Probability of Dose Level being Selected as Optimal for Toxicity and Efficacy
No Dose Level is chosen	0%	9.9%	0.23
1	0%	0%	0
2	0.3%	25.2%	0
3	12.7%	30.0%	0.07
4	86.2%	25.1%	0.69
5	0.8%	9.6%	0.01
6	0%	0.2%	0

Table 20 Results for Accuracy of Dose Selection in the Extended TEQR and mTPI Designs for Monotonically Increasing True DLT Rates and Plateauing True Response Rates: Stopping Rules of the TEQR Design

TEQR Design

Dose Level	% of Times Dose Level is Chosen for Toxicity	% of Times Dose Level is Chosen for Efficacy	Probability of Dose Level being Selected as Optimal for Toxicity and Efficacy
No Dose Level is chosen	0%	11.5%	0.29
1	0.8%	0%	0
2	3.2%	28.0%	0.01
3	31.8%	27.5%	0.18
4	63.7%	26.5%	0.52
5	0.5%	6.3%	0
6	0%	0.2%	0

mTPI Design

Dose Level	% of Times Dose Level is Chosen for Toxicity	% of Times Dose Level is Chosen for Efficacy	Probability of Dose Level being Selected as Optimal for Toxicity and Efficacy
No Dose Level is chosen	0%	11.2%	0.22
1	0%	0%	0
2	0.2%	25.9%	0
3	14.7%	29.1%	0.08
4	85.0%	27.0%	0.70
5	0.1%	6.8%	0
6	0%	0%	0

Table 21 Results for Accuracy of Dose Selection in the Extended TEQR and mTPI Designs for Monotonically Increasing True DLT Rates and Plateauing True Response Rates: Stopping Rules of the mTPI Design

TEQR Design

Dose Level	% of Times Dose Level is Chosen for Toxicity	% of Times Dose Level is Chosen for Efficacy	Probability of Dose Level being Selected as Optimal for Toxicity and Efficacy
No Dose Level is chosen	0%	11.3%	0.28
1	0.6%	0%	0
2	2.6%	25.1%	0.01
3	31.3%	32.3%	0.18
4	64.5%	23.4%	0.52
5	1.0%	7.6%	0.01
6	0%	0.3%	0

mTPI Design

Dose Level	% of Times Dose Level is Chosen for Toxicity	% of Times Dose Level is Chosen for Efficacy	Probability of Dose Level being Selected as Optimal for Toxicity and Efficacy
No Dose Level is chosen	0%	12.1%	0.24
1	0%	0%	0
2	0.3%	25.1%	0
3	12.7%	29.9%	0.07
4	86.2%	24.3%	0.68
5	0.8%	8.6%	0.01
6	0%	0%	0

Table 22 Results for Accuracy of Dose Selection in the Extended TEQR and mTPI Designs for Monotonically Increasing True DLT Rates and Umbrella-Shaped True Response Rates: Stopping Rules of the TEQR Design

TEQR Design

Dose Level	% of Times Dose Level is Chosen for Toxicity	% of Times Dose Level is Chosen for Efficacy (λ=0)	% of Times Dose Level is Chosen for Efficacy (λ=0.01)	% of Times Dose Level is Chosen for Efficacy (λ=0.025)	% of Times Dose Level is Chosen for Efficacy (λ=0.05)	% of Times Dose Level is Chosen for Efficacy (λ=0.1)	% of Times Dose Level is Chosen for Efficacy (λ=0.5)	% of Times Dose Level is Chosen for Efficacy (λ=1)
1	0.8%	0%	0%	0%	0%	0%	0%	0%
2	3.2%	47.1%	34.3%	33.3%	32.4%	31.1%	23.6%	23.5%
3	31.8%	32.1%	29.3%	27.8%	27.0%	25.5%	19.7%	19.7%
4	63.7%	14.5%	22.1%	21.0%	19.8%	17.2%	14.6%	14.6%
5	0.5%	0.3%	0.1%	1.0%	1.0%	1.0%	1.1%	1.1%
6	0%	0%	0%	0%	0%	0%	0%	0%

mTPI Design

Dose Level	% of Times Dose Level is Chosen for Toxicity	% of Times Dose Level is Chosen for Efficacy (λ=0)	% of Times Dose Level is Chosen for Efficacy (λ=0.01)	% of Times Dose Level is Chosen for Efficacy (λ=0.025)	% of Times Dose Level is Chosen for Efficacy (λ=0.05)	% of Times Dose Level is Chosen for Efficacy (λ=0.1)	% of Times Dose Level is Chosen for Efficacy (λ=0.5)	% of Times Dose Level is Chosen for Efficacy (λ=1)
1	0%	0%	0%	0%	0%	0%	0%	0%
2	0.2%	45.2%	31.0%	30.5%	30.1%	29.3%	22.4%	22.3%
3	14.7%	34.9%	31.1%	30.3%	29.6%	28.2%	22.5%	22.5%
4	85.0%	17.2%	24.5%	22.8%	21.1%	18.3%	15.2%	15.2%
5	0.1%	0.5%	1.6%	1.6%	1.6%	1.7%	2%	2%
6	0%	0%	0%	0%	0%	0%	0%	0%

Table 23 Results for Accuracy of Dose Selection in the Extended TEQR and mTPI Designs for Monotonically Increasing True DLT Rates and Umbrella-Shaped True Response Rates: Stopping Rules of the mTPI Design

TEQR Design

Dose Level	% of Times Dose Level is Chosen for Toxicity	% of Times Dose Level is Chosen for Efficacy (λ=0)	% of Times Dose Level is Chosen for Efficacy (λ=0.025)	% of Times Dose Level is Chosen for Efficacy (λ=0.05)	% of Times Dose Level is Chosen for Efficacy (λ=0.1)	% of Times Dose Level is Chosen for Efficacy (λ=0.5)	% of Times Dose Level is Chosen for Efficacy (λ=1)
1	0.6%	0%	0%	0%	0%	0%	0%
2	2.6%	41.7%	29.5%	28.7%	27.3%	20.8%	20.8%
3	31.3%	38.0%	33.8%	32.0%	29.7%	22.6%	22.6%
4	64.5%	14.1%	20.7%	18.7%	16.7%	13.5%	13.5%
5	1.0%	0.4%	1.5%	1.5%	1.5%	1.5%	1.5%

mTPI Design

Dose Level	% of Times Dose Level is Chosen for Toxicity	% of Times Dose Level is Chosen for Efficacy (λ=0)	% of Times Dose Level is Chosen for Efficacy (λ=0.025)	% of Times Dose Level is Chosen for Efficacy (λ=0.05)	% of Times Dose Level is Chosen for Efficacy (λ=0.1)	% of Times Dose Level is Chosen for Efficacy (λ=0.5)	% of Times Dose Level is Chosen for Efficacy (λ=1)
1	0%	0%	0%	0%	0%	0%	0%
2	0.3%	44.6%	30.3%	29.6%	29.1%	21.4%	21.3%
3	12.7%	36.6%	31.9%	31.0%	29.5%	23.3%	23.3%
4	86.2%	15%	22.4%	21.1%	18.9%	16%	16%
5	0.8%	0.5%	1.8%	1.8%	1.9%	2%	2%
6	0%	0%	0%	0%	0%	0%	0%

Table 24 Results for Accuracy of Optimal Dose Selection for Various Designs

Scenario		Dose Level						
		1	2	3	4	5	6	None
1								
	True toxicity, true efficacy rate	.01, .1	.02, .30	.06, .4	.2, .45	.55, 0.55	.89, .6	-
	Trade-off Value Eff-Tox Design	**-0.82**	**-0.43**	**-0.30**	**-0.43**	**-0.77**	**-1.2**	**-**
	% selected by Eff-Tox Design	**0**	**1**	**41**	**53***	**2**	**0**	**3**
	% selected by OBD Design	**8**	**21.6**	**25.3**	**36.7**	**8.3**	**0**	**0**
	% selected by Extended mTPI Design	**0**	**0.08**	**7.01**	**69.22**	**0.72**	**0**	**23**
	% selected by Extended TEQR Design	**0**	**0.66**	**18.06**	**52.7**	**0.91**	**0**	**27.68**
2								
	True toxicity, true efficacy rate	.01, .1	.02, .30	.06, .4	.2, .45	.55, 0.45	.89, .45	-
	Trade-off Value Eff-Tox Design	**-0.82**	**-0.43**	**-0.30**	**-0.43**	**-0.98**	**-1.51**	**-**
	% selected by Eff-Tox Design	**0**	**1**	**42**	**50**	**2**	**0**	**5**
	% selected by OBD Design	**8**	**21.9**	**27.4**	**36.8**	**5.9**	**0**	**0**
	% selected by Extended mTPI Design	**0**	**0.08**	**6.99**	**68.36**	**0.7**	**0**	**23.88**
	% selected by Extended TEQR Design	**0**	**0.65**	**17.97**	**52.12**	**0.88**	**0**	**28.37**
3								
	True toxicity, true efficacy rate	.05, .10	.10, .30	.15, .60	.30, .62	.45, .65		-
	Trade-off Value Eff-Tox Design	**-0.88**	**-0.57**	**-0.04**	**-0.24**	**-0.41**		**-**
	% selected by Eff-Tox Design	**0**	**1**	**47**	**36**	**15**		**0**
	% selected by OBD Isotonic Design	**7.6**	**4.8**	**37.8**	**37.1**	**12.8**		**0**
	% selected by Extended mTPI Design	**0**	**3.6**	**49**	**31.5**	**1.2**		**14.6**
	% selected by Extended TEQR Design	**0**	**6**	**41.5**	**21.9**	**1.5**		**29.2**
4								
	True toxicity, true efficacy rate	.20, .60	.40, .62	.55, .65	.70, .70	.85, .75		-
	Trade-off Value Eff-Tox Design	**-0.12**	**-0.40**	**-0.57**	**-0.70**	**-0.83**		**-**
	% selected by Eff-Tox Design	**66**	**26**	**3**	**0**	**0**		**5**

	% selected by OBD Isotonic Design	**46.1**	**41.4**	**11.3**	**1.2**	**0**		**0**
	% selected by Extended mTPI Design	**69**	**7.5**	**0**	**0**	**0**		**23.5**
	% selected by Extended TEQR Design	**57.6**	**12**	**0.1**	**0**	**0**		**30.3**
5								
	True toxicity, true efficacy rate	.01, .1	.02, .35	.06, .5	.2, .3	.55, 0.2	.89, .05	-
	Trade-off Value Eff-Tox Design	**-0.82**	**-0.33**	**-0.1**	**-0.73**	**-1.48**	**-2.32**	**-**
	% selected by Eff-Tox Design	**1**	**1**	**23**	**11**	**5**	**2**	**57**
	% selected by OBD Isotonic Design	**5.4**	**19.1**	**65.5**	**9.5**	**0.4**	**0**	**0**
	% selected by Extended mTPI Design	**0**	**18.7**	**65.6**	**3.1**	**0**	**0**	**12.6**
	% selected by Extended TEQR Design	**0**	**17.7**	**64.1**	**3.9**	**0**	**0**	**14.3**

*** The dose level marked in green is the dose that is selected most frequently as the optimal dose by each design. The TEQR and mTPI designs here use the standard stopping rules of the mTPI design: maximum sample is reached or dose level 1 is too toxic. All the designs use a cohort size of 5 and maximum sample size of 50.**

Figure 10 Schematic to Depict the Concept of Optimal Dose Selection for Various Dose-Response Curves for the Extended mTPI and TEQR Designs

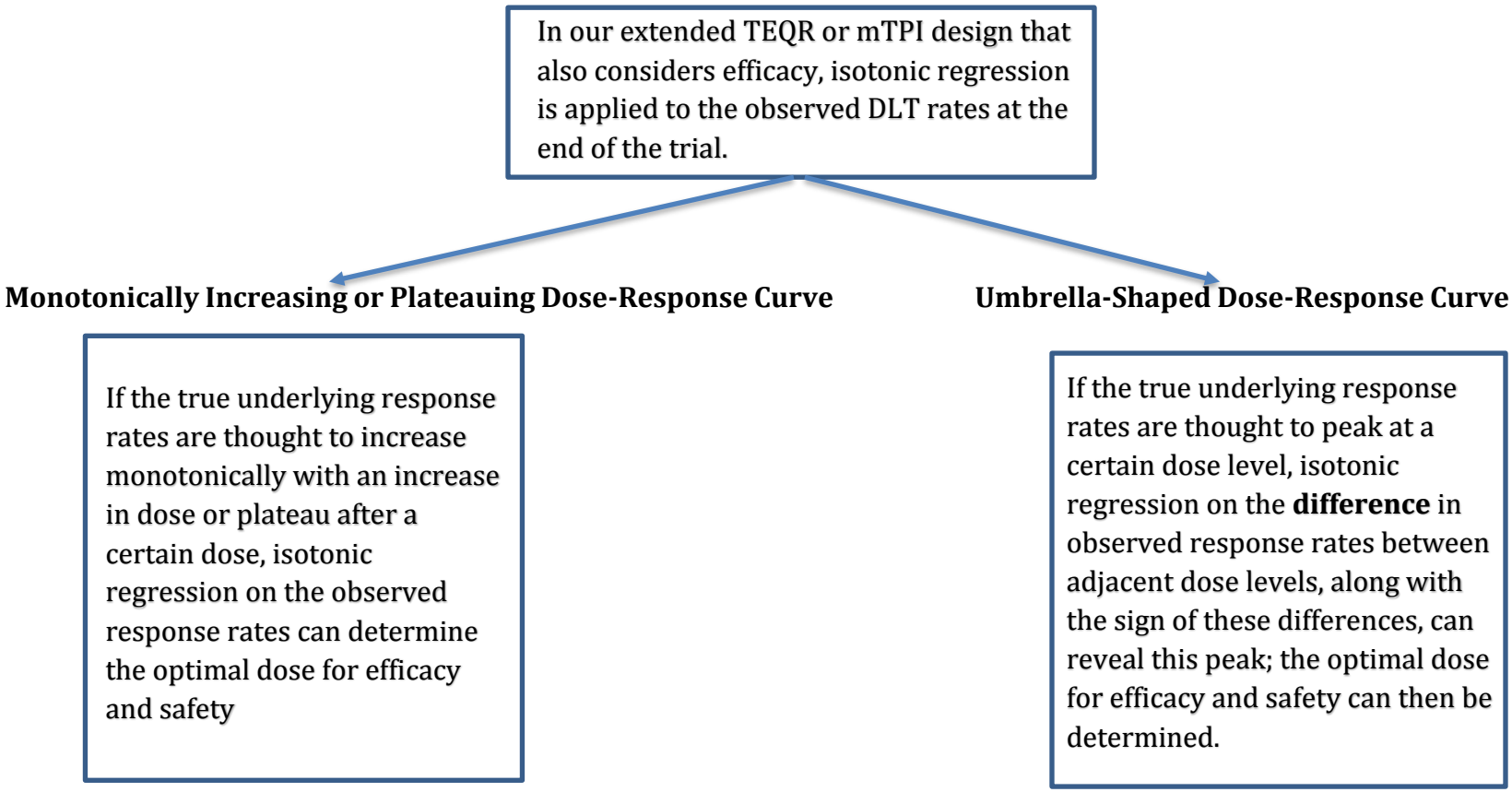


Figure Legend: Schematic of Analysis Method for Different Dose-Response Curves

Figure 11 Schematic to Depict Optimal Dose Selection in the Extended mTPI and TEQR Designs for Dose-Response Curves that Peak at Various Dose Levels

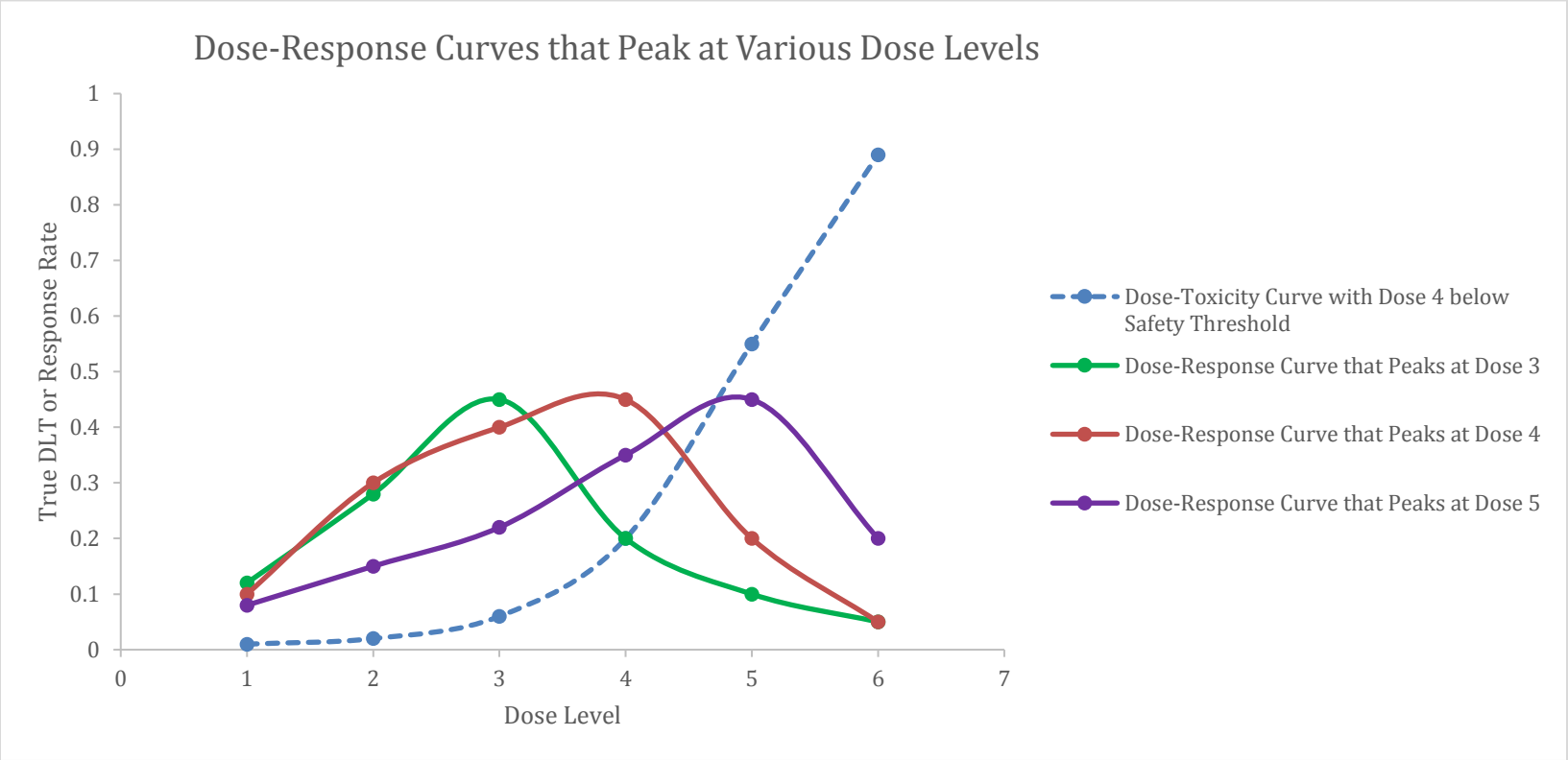


Figure Legend: In this example, Dose level 4 is below the target toxicity rate of 0.33 (blue curve with dashes). For the green dose-response curve with the peak response rate at dose level 3, dose level 3 is chosen as the optimal dose for toxicity and efficacy, assuming the peak response rate is above the efficacy threshold at dose level 3. For the brown dose-response curve with the peak response rate at dose level 4, dose level 4 is chosen as the optimal dose, assuming the peak response rate is above the efficacy threshold at dose level 4. For the purple dose-response curve with the peak response rate at dose level 5, dose level 4 is chosen as the optimal dose, only if the response rate at dose level 4 reaches the efficacy threshold – if not, no dose is chosen as the optimal dose.

Appendix Chapter 4

1. Input parameters for the mTPI and TEQR Designs

Appendix Table 1: Parameters for the mTPI and TEQR Designs:

Parameter	mTPI Design	TEQR Design
Number of simulations	2000	2000
Target toxicity probability p_T	0.2	0.2
ε_1	0.05	0.05
ε_2	0.05	0.05
Starting dose	Dose level 2	Dose level 2
DLT probability deemed to be too toxic to allow further study at that dose level	NA	0.34
Desired sample size at MTD	NA	50
Maximum number of cohorts	NA	30
True DLT rate at each dose level	Values from Tables 15	Values from Table 15
True Response rate at each dose level	Values from Tables 15	Values from Table 15

We start from Dose Level 2 to allow for immediate de-escalation to dose level 1, if required.

2. Results for Dose Selection for the Extended TEQR Design with a Non-Zero Correlation Coefficient between the True Toxicity and Efficacy Rates

The monotonically increasing true DLT and response rates with an increase in dose shown in Table 15 are used in the simulations and isotonic regression is applied independently to the observed DLT and to the observed response rates at the end of each simulated trial.

We use the usual stopping rules of the TEQR design, namely stop the trial when the desired MTD sample size is reached or when dose level 1 is too toxic, with **a non-zero correlation coefficient** between true toxicity and efficacy rates. The results in Appendix Table 2 are based on a sample size at the MTD of 50, a cohort size of 5 and a correlation coefficient r=0.22.

Appendix Table 2: Results for Dose Selection for the Extended TEQR Design for a Non-Zero Correlation between Toxicity and Efficacy
TEQR Design

Dose Level	% of Times Dose Level is Chosen for Toxicity	% of Times Dose Level is Chosen for Efficacy	% of Times Dose Level is Chosen as Optimal for Toxicity and Efficacy
No dose is chosen	0%	5%	28%*
1	0.8%	0%	0%
2	3.2%	28.4%	0.6%
3	31.8%	29.7%	17.1%
4	63.7%	29.0%	53.8%
5	0.5%	7.9%	0.5%

* These results for the % of times each dose level is selected as optimal for toxicity and efficacy are based on simulations and are not analytic calculations based on the % of times each dose level is chosen for toxicity and efficacy, since the correlation coefficient r is not 0.

We choose dose level 4 as the optimal dose for safety and efficacy most frequently in this case. The results for optimal dose selection are similar to those obtained for the "no correlation between the efficacy and toxicity rates" case i.e. r=0 case, with the other input parameters and stopping rules remaining the same.

3. Incorporating Safety and Efficacy: True DLT and Response Rates that Increase Monotonically with an Increase in Dose

We use the usual stopping rules of the TEQR design, namely stop the trial when the desired MTD sample size is reached or when dose level 1 is too toxic. The results in Appendix Table 3 a and b are based on a sample size at the MTD of 50, a cohort size of 5 and a correlation coefficient r=0.

The monotonically increasing true DLT rates shown in Table 15 are used to produce the simulations in Appendix Tables 3 a and 3 b. The response rates with an increase in dose shown in Table 15 are used in the simulations in Appendix Table 3 a. Isotonic regression is applied to the observed DLT rates and isotonic regression is applied to the **difference** in observed response rates between adjacent dose levels at the end of each simulated trial (Appendix Tables 3 a and 3 b).

We calculate the difference in the observed response rates between adjacent dose levels at the end of each simulated trial i.e. between dose level 1 and 2, 2 and 3, 3 and 4 etc. We then investigate where the differences in response rates between adjacent dose levels switch from a negative sign to a positive sign after applying isotonic regression to these differences, to determine whether there is a peak at a certain dose level in the observed response rates. If in a simulation the difference in response rates between dose levels 2 and 3 is negative and that between dose levels 3 and 4 is positive, then dose level 3, is considered the peak. If the difference in response rates between dose levels 1 and 2 is negative, that between dose levels 2 and 3 is zero and that between dose levels 3 and 4 is positive, then dose level 3, is the peak. If all the differences between adjacent dose levels are negative in a simulation, then no dose level chosen as the peak for that simulation. If all the differences between adjacent dose levels are positive in a simulation, then dose level 1 is the peak.

In this example, dose level 3 is what is picked most frequently as the peak but it is

selected only 23% of the time. Also, note that no dose level is chosen as the peak dose level for efficacy 44% of the time. Thus, there is no peak in the observed response rates at dose level 3. Based on the true response rates in Table 15, the difference in true response rates between adjacent dose levels is always negative and there is no peak at dose level 3. Hence, the results in Appendix Table 3 a reflect this underlying scenario.

Appendix Table 3 a: Results for Accuracy of Dose Selection for the Extended TEQR Design for Monotonically Increasing True DLT and Response Rates Using the Difference in Observed Response Rates

TEQR Design

Dose Level	% of Times Dose Level is Chosen for Toxicity	% of Times Dose Level is a Peak for Efficacy
No dose level is chosen	0%	44.2%
1	0.8%	0.1%
2	3.2%	12%
3	31.8%	22.9%
4	63.7%	12.8%
5	0.5%	7.9%
6	0%	0.1%

As another example, if the true response rate at dose level 1, 2, 3, 4, 5 and 6 are 0.1, 0.2, 0.3, 0.4. 0.5 and 0.6 respectively i.e. the difference in true response rates between adjacent dose levels is always -0.1, and isotonic regression is applied to the observed differences at the end of each of the 1000 simulations, the results in Appendix Table 3 b are obtained. It is seen, in this scenario, that dose level 3 is picked as the peak most

frequently, but only 16% of the time, and that no dose level is selected as the peak 53% of the time, reflecting the underlying scenario in true response rates where there is no peak.

Appendix Table 3 b: Results for Accuracy of Dose Selection for the Extended TEQR Design for Monotonically Increasing True DLT and Response Rates Using the Difference in Observed Response Rates

TEQR Design

Dose Level	% of Times Dose Level is Chosen for Toxicity	% of Times Dose Level is a Peak for Efficacy
No dose level is chosen	0%	53.0%
1	0.8%	0.1%
2	3.2%	8.2%
3	31.8%	15.5%
4	63.7%	15.2%
5	0.5%	7.9%
6	0%	0.1%

Hence, in general, when the true response rates are monotonically increasing with an increase in dose, applying isotonic regression on the difference in response rates between adjacent dose levels will show that no dose level is selected as the peak dose most frequently; thus, in this case, applying isotonic regression on the observed response rates themselves can help determine the dose to be chosen for efficacy.

4. Incorporating Safety and Efficacy: Plateauing True Response Rates with an Increase in Dose

We use the usual stopping rules of the TEQR design, namely stop the trial when the desired MTD sample size is reached or when dose level 1 is too toxic. The results in Appendix Table 4 a and b are based on a sample size at the MTD of 50, a cohort size of 5 and a correlation coefficient r=0.

The monotonically increasing true DLT rates shown in Table 15 are used to produce the simulations in Appendix Tables 4 a and 4 b. The plateauing true response rates with an increase in dose shown in Table 15 are used in the simulations in Appendix Table 4 a. Isotonic regression is applied to the observed DLT rates and isotonic regression is applied to the **difference** in observed response rates between adjacent dose levels at the end of each simulated trial (Appendix Tables 4 a and 4 b).

We calculate the difference in the observed response rates between adjacent dose levels at the end of each simulated trial and apply isotonic to these differences. We then investigate where these differences switch signs after isotonic regression, to help determine whether there is a peak at a certain dose level in the observed response rates. In this example, there is no peak 36% of the time, dose level 3 is chosen as the peak most frequently, and dose level 4 is chosen as the peak only 18% of the time. Thus, there is no clear peak or plateau in response rates at dose level 4.

Based on the true underlying response rates in Table 15, there is a plateau from dose level 4. However, the results in Appendix Table 4 a do not clearly show that there is a peak or plateau in response rates at dose level 4.

Appendix Table 4 a: Results for Accuracy of Dose Selection for the Extended TEQR Design for Monotonically Increasing True DLT Rates and Plateauing Response Rates Using the Difference in Observed Response Rates

TEQR Design

Dose Level	% of Times Dose Level is Chosen for Toxicity	% of Times Dose Level is a Peak for Efficacy
No dose level is chosen	0%	36.2%
1	0.8%	0.1%
2	3.2%	13.5%
3	31.8%	24.7%
4	63.7%	18.3%
5	0.5%	7.0%
6	0%	0.2%

As another example, if the true response rate at dose level 1, 2, 3, 4, 5 and 6 are 0.1, 0.35, 0.5, 0,5, 0.5 and 0.5 respectively, and isotonic regression is applied to the observed differences at the end of each of the 1000 simulations, the results in Appendix Table 4 b show that dose level 3, the dose level that is selected as the peak most frequently, is selected only 34% of the time. However, no dose level is selected as the peak almost as frequently, at 32% of the time. Thus, the results in Appendix Table 4 b do not clearly show that there is a peak or plateau in response rates at dose level 3, as present in the true underlying response rates.

Appendix Table 4 b: Results for Accuracy of Dose Selection for the Extended TEQR Design for Monotonically Increasing True DLT Rates and Plateauing Response Rates Using the Difference in Observed Response Rates

TEQR Design

Dose Level	% of Times Dose Level is Chosen for Toxicity	% of Times Dose Level is a Peak for Efficacy
No dose level is chosen	0%	31.8%
1	0.8%	0%
2	3.2%	11.7%
3	31.8%	34.2%
4	63.7%	13.6%
5	0.5%	8.7%

Hence, applying isotonic regression on the observed response rates themselves in the case the response rates plateau with an increase in dose will show that no dose level is selected as the peak dose frequently; thus, in this case, when there is no clear peak, applying isotonic regression on the observed response rates themselves can help determine the dose to be chosen for efficacy.

5. Incorporating Safety and Efficacy: Response Rates that Follow an Umbrella-Shaped Dose-Response Curve

We use the usual stopping rules of the TEQR design, namely stop the trial when the desired MTD sample size is reached or when dose level 1 is too toxic. The results in

Appendix Table 5 a, b and c are based on a sample size at the MTD of 50, a cohort size of 5 and a correlation coefficient r=0.

The monotonically increasing true DLT rates with an increase in dose shown in Table 15 are used to produce the simulations in Appendix Tables 5 a, 5 b and 5 c. The umbrella-shaped true response rates shown in Table 15, where the response rate peaks at dose level 4, are used in the simulations in Appendix Table 5 a. Isotonic regression is applied to the observed DLT rates and isotonic regression is applied to the **difference** in observed response rates between adjacent dose levels at the end of each simulated trial (Appendix Tables 5 a, 5 b and 5 c).

We calculate the difference in the observed response rates between adjacent dose levels at the end of each simulated trial and apply isotonic to these differences. We then investigate where these differences switch signs after isotonic regression, to help determine whether there is a peak at a certain dose level in the observed response rates. In this example, dose level 3 is chosen slightly more frequently than dose level 4 as the peak. Based on the true underlying response rates in Table 15, there is a peak at dose level 4. Thus, the results in in Appendix Table 5 a do not reflect this peak at dose level 4 as clearly. It is seen from the examples below that when the peak in response rates is at a lower dose level, such as at dose level 3, it is revealed clearly by this method. In other words, not as many patients are dosed at the higher

dose levels compared to at the lower doses, and it is difficult to reveal the peak in response rates when it is at these higher dose levels.

Appendix Table 5 a: Results for Accuracy of Dose Selection for the Extended TEQR Design for Monotonically Increasing True DLT Rates and Umbrella-Shaped Response Rates Using the Difference in Observed Response Rates

TEQR Design

Dose Level	% of Times Dose Level is Chosen for Toxicity	% of Times Dose Level is a Peak for Efficacy
No dose level is chosen	0%	22.7%
1	0.8%	0.1%
2	3.2%	15.5%
3	31.8%	32.0%
4	63.7%	27.6%
5	0.5%	2.1%

If the true response rate at dose level 1, 2, 3, 4, 5 and 6 are 0.1, 0.35, 0.5, 0,3, 0.2 and 0.05 respectively, and isotonic regression is applied to the observed differences at the end of each of the 1000 simulations, the results in Appendix Table 5 b show that dose level 3 is chosen as the peak most frequently, consistent with the peak at dose level 3 in the true underlying response rates. Dose level 3 is selected as the optimal dose 63% of the time (Figure 11 explains how the optimal dose is selected at the end of each simulation for a dose-response curve with a peak).

Appendix Table 5 b: Results for Accuracy of Dose Selection for the Extended TEQR Design for Monotonically Increasing True DLT Rates and Umbrella-Shaped Response Rates Using the Difference in Observed Response Rates

TEQR Design

Dose Level	% of Times Dose Level is Chosen for Toxicity	% of Times Dose Level is a Peak for Efficacy	% of Times Dose Level is Selected as Optimal for Safety and Efficacy
No dose level is chosen	0%	9%	15.8%
1	0.8%	0%	0%
2	3.2%	20%	19.4%
3	31.8%	62.7%	62.7%
4	63.7%	5.1%	2%
5	0.5%	3.2%	0.1%

As a final example, if the true response rate at dose level 1, 2, 3, 4, 5 and 6 are 0.05, 0.1, 0.45, 0,3, 0.2, 0.05 respectively, and isotonic regression is applied to the observed differences at the end of each of the 1000 simulations, the results in Appendix Table 5 c show that dose level 3 is chosen as the peak most frequently, consistent with the peak at dose level 3 in the true underlying response rates. Dose level 3 is selected as the optimal dose 63% of the time.

Appendix Table 5 c: Results for Accuracy of Dose Selection for the Extended TEQR Design for Monotonically Increasing True DLT Rates and Umbrella-Shaped Response Rates Using the Difference in Observed Response Rates

TEQR Design

Dose Level	% of Times Dose Level is Chosen for Toxicity	% of Times Dose Level is a Peak for Efficacy	% of Times Dose Level is Selected as Optimal for Safety and Efficacy
No dose level is chosen	0%	15.0%	32%
1	0.8%	0.3%	0%
2	3.2%	1.7%	1.4%
3	31.8%	68.7%	63.4%
4	63.7%	9.4%	3.1%
5	0.5%	4.9%	0.1%

Thus, if the true underlying response rates have a clear peak at a certain dose level (umbrella-shaped dose-response curve), isotonic regression on the difference in observed response rates between adjacent dose levels, along with the sign of these differences, can be used to reveal this peak, and determine an optimal dose for toxicity and efficacy (Figure 11). The technique works well when the peak in response rates is at a lower dose level rather than a higher dose level where few patients may be dosed.

6. Input Parameters Used in Simulations for the Eff-Tox and OBD Isotonic Designs

Input parameters Used in the Eff-Tox Package for the Eff-Tox Design

Probability of Toxicity and Efficacy Limits for Dose Acceptability Rules

Parameter	Value
Prob(tox) upper limit (π_T*)	0.33000
Lower prob cutoff for prob of toxicity ($p_{T,L}$)	0.10000
Prob(eff) lower limit (π_E*)	0.40000
Lower prob cutoff for prob of efficacy ($p_{E,L}$)	0.10000

Trade-off Function Elicited Points (3 points to define the trade-off function contour)

	π_E	π_T
($\pi_{1,E}$*, 0)	0.50000	0.00000
(1, $\pi_{2,T}$*)	1.00000	0.65000
($\pi_{3,E}$, $\pi_{3,T}$)	0.70000	0.25000

Elicited Means (Prior Toxicity, Prior Efficacy)

Dose	Toxicity	Efficacy
1	0.0200	0.2000
2	0.0400	0.4000
3	0.0600	0.6000
4	0.0800	0.8000
5	0.1000	0.9000

Input parameters Used in the OBD Isotonic Design

Cohort size = 5

Number of cohorts = 10

phi = upper bound of toxicity rate = 0.33

ct = threshold for posterior probability of toxicity (any dose with toxicity probability larger than ct is excluded from the admissible set of doses) = 0.9

Number of simulations = 10000

The true DLT and response rates at each dose level are as given in Table 24.

The R code given at the following URL was used along with the input parameters given above:

http://odin.mdacc.tmc.edu/~yyuan/Software/TargetAgent/targetAgentDF.r

DISCUSSION AND CONCLUSIONS

In this thesis, we first systematically compared the statistical operating characteristics of 11 existing Phase 1 rule-based and model-based oncology dose-finding designs that target or pre-specify a DLT rate of ~0.2, for three sets of true underlying DLT rates. These DLT rates were generated from a specific logistic, log-logistic and linear dose-toxicity curve at the same dosages. We found that all the designs examined select the MTD much more accurately when there is a clear separation between the true DLT rate at the MTD and the rates at the dose level immediately above and below it, such as for the DLT rates generated using the chosen logistic dose-toxicity curve. Not surprisingly, we found that the 3+3 dose-finding design under-doses a large percentage of patients and is not very accurate in selecting the MTD in all the cases considered. Among the rule-based designs studied, the 5+5 a design picked the MTD as accurately as the model based designs for the true DLT rates generated using the chosen log-logistic and linear dose-toxicity curves, but required enrolling a higher number of patients than the other designs. The model based designs, mTPI, TEQR, BOIN, CRM and EWOC designs, performed well on the whole – they assigned the maximum percentage of patients to the MTD, and also had a reasonably high probability of picking the right dose level as the MTD across the three dose toxicity curves examined. We also found that it is critical to pick a design that is aligned with the true DLT rate of interest. Thus, the simulation results contained in this thesis provide considerable information on design property trade-offs, and the means to explore additional settings. However, this is not an exhaustive comparison

of all the current Phase 1 oncology designs. We have covered multiple commonly used ones but future comparisons or studies could include other dose-finding designs such as the Time-to-Event CRM (TITE-CRM) design, the Rolling 6 design, the Bayesian Logistic Regression Method (BLRM) and the recently revised mTPI design, called the mTPI-2 design, proposed by Ji et al..

However, the sample size of these Phase I oncology trials is very limited and it is difficult to accurately predict the MTD with such a small number of patients. Hence, we next studied the effect of sample size and cohort size on the accuracy of dose selection in early phase oncology designs, finding that an adequate sample size is crucial. We then proposed a new design with a larger sample size that encompasses the objectives of both safety and efficacy and that is simpler to implement than the existing Phase 1/2 seamless designs. In particular, we proposed the 20+20 accelerated titration design, a simple rule-based integrated Phase 1/2 trial design that selects an optimal dose for toxicity and efficacy via Bayesian decision rules at the end. Our simulations of the 20+20 accelerated titration design yielded a mean sample size of ~42 patients for the chosen true underlying DLT and response rates and stopping rules and showed that with this sample size, the design can robustly pick a dose that is optimal for both efficacy and safety. We showed via simulated examples that it performed as well as or better than the Eff-Tox design and the Optimal Biological Dose Isotonic design for the scenarios considered. It also performed better than a 3+3 Phase 1 design followed by a Phase 2 design. Further, this technique used

in the 20+20 accelerated titration design of selecting an optimal dose for safety and efficacy via Bayesian decision rules at the end of the trial could also be applied more generally to other Phase 1 designs in future work.

We also extended the TEQR and mTPI dose-finding oncology designs to choose an optimal dose for both safety and efficacy by considering correlated Bernoulli distributions for the true underlying toxicity and efficacy rates. In our simulations, we assumed that the true DLT rates increase monotonically with an increase in dose but did not assume that the true response rates increase monotonically with an increase in dose; we allow multiple types of curves for dose response (monotonically increasing, plateau, or umbrella-shaped). In this context, we applied isotonic regression to determine a dose that is optimal for both safety and efficacy. We showed that the extended TEQR and mTPI designs performed as well as or better than the Eff-Tox design and the Optimal Biological Dose Isotonic design in terms of accuracy of optimal dose selection for the scenarios considered.

Several further extensions to this thesis work could be considered. Future work could consider:

1) extending these designs to systematically and rigorously implement rules in their dose finding algorithms for late occurring DLTs, i.e. DLTs occurring outside the protocol defined observation period of (say) the first 2 cycles of the study drug. It would be useful to extend some of these designs such as the TEQR design to

account for DLTs occurring outside this window, similar to the TITE-CRM design where the timing of late-onset DLTs is incorporated [116, 117]. In general, it would be useful to extend the TEQR and other designs to account for the time that the patient is on study before having a DLT for patients who experience a DLT, and also to account for patients who do not experience a DLT (see reference [118]).

2) extending these designs to incorporate PK and/or biomarker data in their dose finding algorithms.
3) extending these dose finding oncology designs to include, in addition to a binary endpoint for safety i.e. DLT vs no DLT, a continuous endpoint for efficacy; for example, we can include a continuous immune response, as opposed to a binary response such as Complete Response/Partial Response (CR/PR) vs no CR/PR, as we have done in the 20+20 acceleration titration design and in extending the mTPI and TEQR designs. We would then have to consider how an optimal dose for safety and efficacy can be selected taking into account the correlation between a binary endpoint for safety and a continuous one for efficacy [119].
4) extending these early phase designs to combination studies of two drugs for both efficacy and toxicity i.e. to determine an optimal dose(s) for efficacy and toxicity for a combination of two drugs. This is a challenging but useful problem to tackle due to the increasing number of drug combination studies being performed in oncology.

APPENDIX CHAPTER FOR PROGRAM CODES USED IN THE THESIS

Codes for Chapter 2

SAS code for all designs where only escalation is allowed (3+3, 2+4, 4+4 a, 4+4 b, 5+5 a, 5+5 b, 3+3+3, 3+1+1):

```
%macro alldesigns(design1, equation1);
data simi1;
call streaminit(1);

array a{1000,10};
array sumi{1000};
array dosel[10000];
array dosemtd{10000};
array doseover{10000};
array doseunder{10000};
array dltmtd{10000};
array dltover{10000};
array dltunder{10000};
array dltoverall{10000};
array mtddltrate{10000};
array mtdpop{10000};
array totalpop{10000};
array totaldlt{10000};
array peopledosel{10};
array dltdosel{10};
sumd=0;
sump=0;



/** program works for 3+3, 2+4, 4+4 a, 4+4 b, 5+5 a, 5+5 b, 3+3+3, or
3+1+1 designs and 3 dose-toxicity curves **/

length design $10;
design="&design1";

length equation $20;
equation="&equation1";

if design='3+3' then do;
no1=3;
no2=4;
no3=6;
o1=   1;
o2=   3;
o3=   1;
o4=   6;
o5=   1;
o6=   6;
```

```
end;
if design='2+4' then do;
no1=2;
no2=3;
no3=6;
o1=   1;
o2=   2;
o3=   1;
o4=   6;
o5=   1;
o6=   6;
end;
if design='4+4a' then do;
no1=4;
no2=5;
no3=8;
o1=   2;
o2=   4;
o3=   2;
o4=   8;
o5=   2;
o6=   8;
end;
if design='4+4b' then do;
no1=4;
no2=5;
no3=8;
o1=   2;
o2=   4;
o3=   3;
o4=   8;
o5=   3;
o6=   8;
end;

if design='3+1+1' then do;
no1=3;
o1=   2;
o2=   3;
o3=   2;
o4=   4;
o5=   2;
o6=   5;
end;
if design='3+3+3' then do;
no1=3;
no2=4;
no3=6;
o1=   1;
o2=   3;
o3=   2;
o4=   6;
o5=   2;
o6=   9;
```

```
end;
if design='5+5a' then do;
no1=5;
no2=6;
no3=10;
o1=   2;
o2=   5;
o3=   2;
o4=   10;
o5=   2;
o6=   10;
end;
if design='5+5b' then do;
no1=5;
no2=6;
no3=10;
o1=   2;
o2=   5;
o3=   3;
o4=   10;
o5=   3;
o6=   10;
end;

/* program can consider one of 3 dose-toxicity curves; change these
parameters if you want to change the dose-toxicity curve */

startdose = 100;
mtddose=334;
dltrstartdose=0.01;
dltrmtd=0.2;
doselevelmtd=3;

if equation='linear' then do;

coeff2=1/(mtddose-startdose)*(dltrmtd-dltrstartdose);
coeff1= dltrstartdose - startdose*coeff2;

end;

if equation='logistic' then do;

coeff2=1/(mtddose-startdose)*( log(dltrmtd/(1-dltrmtd))-
log(dltrstartdose/(1-dltrstartdose)) );
coeff1= log(dltrstartdose/(1-dltrstartdose)) - startdose*coeff2;

end;

if equation='loglogistic' then do;

coeff2=1/(log(mtddose)-log(startdose))*( log(dltrmtd/(1-dltrmtd))-
log(dltrstartdose/(1-dltrstartdose)) );
coeff1= log(dltrstartdose/(1-dltrstartdose)) - log(startdose)*coeff2;
```

```
end;



do k=1 to 10000;


do i1=1 to 1000;
do j1=1 to 10;
a[i1,j1]=.;
sumi[i1]=.;
end;
end;

i=0;
sumi[1]=0;
peoplec=0;
atmtd=0;
belowmtd=0;
abovemtd=0;
dltc=0;
dltatmtd=0;
dltbelowmtd=0;
dltabovemtd=0;

dose=100;

/** escalate until the number of DLTs exceeds what is allowed by the
design **/
do until ((sumi[i]>o1 and sumins=o2) or (sumi[i]>o3 and sumins=o4) or
(sumi[i]>o5 and sumins=o6));
    i = i+1;
      sumi[i]=0;
      sumins=0;


if i=1 then frac=1;
if i=2 then frac=2;
if i=3 then frac=1.67;
if i=4 then frac=1.5;
if i=5 then frac=1.4;
if i=6 then frac=1.33;
if i>=7 then frac=1.33;
dose=dose*frac;

if equation='linear' then dr=min(coeff1+coeff2*dose,1);

if equation='logistic' then dr=exp(coeff1+coeff2*dose)/(1+
exp(coeff1+coeff2*dose));

if equation='loglogistic' then dr=exp(coeff1+coeff2*log(dose))/(1+
exp(coeff1+coeff2*log(dose)));

      do j=1 to no1;
```

```
            a[i,j]=rand('Bernoulli', dr);
            sumi[i]=sumi[i]+a[i,j];
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            sumins=sumins+1;
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=sum(atmtd,1);
            if i>doselevelmtd then abovemtd=sum(abovemtd,1);
            if i<doselevelmtd then belowmtd=sum(belowmtd,1);
            dltc=dltc+a[i,j];
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
      end;

if (design='3+3' or design='2+4' or design='3+3+3') then do;
if sumi[i]=1 then do;
do j=no2 to no3;
            a[i,j]=rand('Bernoulli', dr);
            sumi[i]=sumi[i]+a[i,j];
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            sumins=sumins+1;
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=sum(atmtd,1);
            if i>doselevelmtd then abovemtd=sum(abovemtd,1);
            if i<doselevelmtd then belowmtd=sum(belowmtd,1);
            dltc=dltc+a[i,j];
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
end;
end;
end;

if (design='4+4a' or design='5+5a') then do;
if (sumi[i]=1 or sumi[i]=2) then do;
do j=no2 to no3;
            a[i,j]=rand('Bernoulli', dr);
            sumi[i]=sumi[i]+a[i,j];
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            sumins=sumins+1;
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=sum(atmtd,1);
            if i>doselevelmtd then abovemtd=sum(abovemtd,1);
            if i<doselevelmtd then belowmtd=sum(belowmtd,1);
            dltc=dltc+a[i,j];
```

```
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
end;
end;
end;

if (design='4+4b' or design='5+5b') then do;
if sumi[i]=2 then do;
do j=no2 to no3;
            a[i,j]=rand('Bernoulli', dr);
            sumi[i]=sumi[i]+a[i,j];
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            sumins=sumins+1;
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=sum(atmtd,1);
            if i>doselevelmtd then abovemtd=sum(abovemtd,1);
            if i<doselevelmtd then belowmtd=sum(belowmtd,1);
            dltc=dltc+a[i,j];
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
end;
end;
end;

if design='3+3+3' then do;
if (sumi[i]=2 and sumins=6) then do;
do j=7 to 9;
            a[i,j]=rand('Bernoulli', dr);
            sumi[i]=sumi[i]+a[i,j];
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            sumins=sumins+1;
            peoplec=peoplec+1;
      if i=doselevelmtd then atmtd=sum(atmtd,1);
            if i>doselevelmtd then abovemtd=sum(abovemtd,1);
            if i<doselevelmtd then belowmtd=sum(belowmtd,1);
            dltc=dltc+a[i,j];
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
end;
end;
end;

if design='3+1+1' then do;
if ((sumi[i]=1) or (sumi[i]=2)) then do;
            a[i,4]=rand('Bernoulli', dr);
```

```
            sumi[i]=sumi[i]+a[i,4];
            sumd=sumd+a[i,4];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            sumins=sumins+1;
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=sum(atmtd,1);
            if i>doselevelmtd then abovemtd=sum(abovemtd,1);
            if i<doselevelmtd then belowmtd=sum(belowmtd,1);
            dltc=dltc+a[i,j];
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
end;

if ((sumi[i]=2) and (sumins=4)) then do;
            a[i,5]=rand('Bernoulli', dr);
            sumi[i]=sumi[i]+a[i,5];
            sumd=sumd+a[i,5];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            sumins=sumins+1;
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=sum(atmtd,1);
            if i>doselevelmtd then abovemtd=sum(abovemtd,1);
            if i<doselevelmtd then belowmtd=sum(belowmtd,1);
            dltc=dltc+a[i,j];
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
end;
end;


end;
dosel[k]=i;
dosemtd[k]=atmtd/peoplec*100;
doseover[k]=abovemtd/peoplec*100;
doseunder[k]=belowmtd/peoplec*100;
dltmtd[k]=dltatmtd/dltc*100;
dltover[k]=dltabovemtd/dltc*100;
dltunder[k]=dltbelowmtd/dltc*100;

dltoverall[k]=dltc/peoplec;
if atmtd ne 0 then mtddltrate[k]=dltatmtd/atmtd;
else mtddltrate[k]=0;
mtdpop[k]=atmtd;
totalpop[k]=peoplec;
totaldlt[k]=dltc;

end;
avgd=sumd/10000;
```

```
avgp=sump/10000;
ave= mean(of dosel1-dosel10000);
dev=STD(of dosel1-dosel10000);
serror=STDERR(of dosel1-dosel10000);
max= max(of dosel1-dosel10000);
median= median(of dosel1-dosel10000);
pctl_25=pctl(25, of dosel1-dosel10000);
pctl_50=pctl(50, of dosel1-dosel10000);
pctl_75=pctl(75, of dosel1-dosel10000);
avgdosemtd=mean(of dosemtd1-dosemtd10000);
avgdoseover=mean(of doseover1-doseover10000);
avgdoseunder=mean(of doseunder1-doseunder10000);
avgdltmtd=mean(of dltmtd1-dltmtd10000);
avgdltover=mean(of dltover1-dltover10000);
avgdltunder=mean(of dltunder1-dltunder10000);
avgdltrate=mean(of dltoverall1-dltoverall10000);
avgmtddltrate=mean(of mtddltrate1-mtddltrate10000);
avgmtdpop=median(of mtdpop1-mtdpop10000);
avgtotalpop=median(of totalpop1-totalpop10000);
avgtotalpop1=mean(of totalpop1-totalpop10000);
avgtotaldlt=median(of totaldlt1-totaldlt10000);
run;

/* print design, highest dose level reached/dose level at which the
trial stops, dose at highest dose level, dlt rate at the highest dose
level, factor to multiply present dose to get the next dose level, avg
(+STD), max and median number of dose levels explored, median and avg
number of patients on trial, median number of DLTs, avg DLT rate of the
trial, avg DLT rate at the MTD, median MTD pop, avg percentage of
patients dosed at MTD, under-dosed and over-dosed, avg number of
patients and DLTs at each dose level */

proc print;
var design i dose dr frac ave dev max median avgtotalpop avgtotalpop1
avgtotaldlt avgdltrate avgmtddltrate avgmtdpop avgdosemtd avgdoseunder
avgdoseover peopledosel1-peopledosel10 dltdosel1-dltdosel10;
run;

data simi2;
set simi1;
array v_[*] dosel1-dosel10000;
array counts[10];
call missing (of counts[*]);
do i = 1 to dim(v_);
     counts[v_[i]] + 1;
end;
if counts1=. then counts1=0;
if counts2=. then counts2=0;
if counts3=. then counts3=0;
if counts4=. then counts4=0;
if counts5=. then counts5=0;
if counts6=. then counts6=0;
if counts7=. then counts7=0;
if counts8=. then counts8=0;
```

```
if counts9=. then counts9=0;
if counts10=. then counts10=0;

run;

/* times out of 10000 that the trial stops at each dose level, from
which the percentage of times each dose level is choosen as the MTD can
be determined - the MTD is one dose level below the dose level at which
the trial stops */
proc print;
var counts1-counts10;
run;

%mend alldesigns;

%alldesigns(3+3, logistic);
```

SAS code for accelerated titration design:

```
data simi1;
call streaminit(1);

array a{1000,6};
array sumi{1000};
array flag{1000};
array dosel[10000];
array dosemtd{10000};
array doseover{10000};
array doseunder{10000};
array dltmtd{10000};
array dltover{10000};
array dltunder{10000};
array dltoverall{10000};
array mtddltrate{10000};
array mtdpop{10000};
array totalpop{10000};
array totaldlt{10000};
array peopledosel{10};
array dltdosel{10};

sumd=0;
sump=0;

/* program can handle one of 3 dose-toxicity equations */
length equation $20;
equation='logistic';

startdose = 100;
mtddose=334;
dltrstartdose=0.01;
dltrmtd=0.2;
doselevelmtd=3;
```

```
if equation='linear' then do;

coeff2=1/(mtddose-startdose)*(dltrmtd-dltrstartdose);
coeff1= dltrstartdose - startdose*coeff2;

end;

if equation='logistic' then do;

coeff2=1/(mtddose-startdose)*( log(dltrmtd/(1-dltrmtd))-
log(dltrstartdose/(1-dltrstartdose)) );
coeff1= log(dltrstartdose/(1-dltrstartdose)) - startdose*coeff2;

end;

if equation='loglogistic' then do;

coeff2=1/(log(mtddose)-log(startdose))*( log(dltrmtd/(1-dltrmtd))-
log(dltrstartdose/(1-dltrstartdose)) );
coeff1= log(dltrstartdose/(1-dltrstartdose)) - log(startdose)*coeff2;

end;


do k=1 to 10000;

do i1=1 to 1000;
do j1=1 to 6;
a[i1,j1]=.;
sumi[i1]=.;
flag[i1]=.;
end;
end;

i=0;
sumi[1]=0;

peoplec=0;
atmtd=0;
belowmtd=0;
abovemtd=0;
dltc=0;
dltatmtd=0;
dltbelowmtd=0;
dltabovemtd=0;

dose=100;

do until (a[i,1]=1);
    i = i+1;
      sumi[i]=0;
```

```
    if i=1 then frac=1;
    if i=2 then frac=2;
    if i=3 then frac=1.67;
    if i=4 then frac=1.5;
    if i=5 then frac=1.4;
    if i=6 then frac=1.33;
    if i>=7 then frac=1.33;

    dose=dose*frac;

    if equation='linear' then dr=min(coeff1+coeff2*dose,1);

    if equation='logistic' then dr=exp(coeff1+coeff2*dose)/(1+
exp(coeff1+coeff2*dose));

    if equation='loglogistic' then dr=exp(coeff1+coeff2*log(dose))/(1+
exp(coeff1+coeff2*log(dose)));

      a[i,1]=rand('Bernoulli', dr);
      sumi[i]=sumi[i]+a[i,1];
      sumd=sumd+a[i,1];
      sump=sump+1;
      peopledosel[i]=sum(peopledosel[i],0.0001);
      dltdosel[i]=sum(dltdosel[i],a[i,1]/10000);
      peoplec=peoplec+1;
      if i=doselevelmtd then atmtd=atmtd+1;
      if i>doselevelmtd then abovemtd=abovemtd+1;
      if i<doselevelmtd then belowmtd=belowmtd+1;
      dltc=dltc+a[i,1];
      if i=doselevelmtd then dltatmtd=dltatmtd+a[i,1];
      if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,1];
      if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,1];
end;

/** switch to 3+3 design once a DLT is observed in a patient **/
do j=2 to 3;
            a[i,j]=rand('Bernoulli', dr);
            sumi[i]=sumi[i]+a[i,j];
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=atmtd+1;
            if i>doselevelmtd then abovemtd=abovemtd+1;
            if i<doselevelmtd then belowmtd=belowmtd+1;
            dltc=dltc+a[i,j];
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
end;


if sumi[i]=1 then do;
```

```
do j=4 to 6;
            a[i,j]=rand('Bernoulli', dr);
            sumi[i]=sumi[i]+a[i,j];
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=atmtd+1;
            if i>doselevelmtd then abovemtd=abovemtd+1;
            if i<doselevelmtd then belowmtd=belowmtd+1;
            dltc=dltc+a[i,j];
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
end;
end;
if sumi[i]=1 then flag[i]=1;
else flag[i]=0;
jj1=i;

if flag[jj1]=1 then do;
do until (sumi[i]>1);
    i = i+1;
      sumi[i]=0;

      if i=1 then frac=1;
    if i=2 then frac=2;
    if i=3 then frac=1.67;
    if i=4 then frac=1.5;
    if i=5 then frac=1.4;
    if i=6 then frac=1.33;
    if i>=7 then frac=1.33;

    dose=dose*frac;

      if equation='linear' then dr=min(coeff1+coeff2*dose,1);

    if equation='logistic' then dr=exp(coeff1+coeff2*dose)/(1+
exp(coeff1+coeff2*dose));

    if equation='loglogistic' then dr=exp(coeff1+coeff2*log(dose))/(1+
exp(coeff1+coeff2*log(dose)));



      do j=1 to 3;
            a[i,j]=rand('Bernoulli', dr);
            sumi[i]=sumi[i]+a[i,j];
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
          dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            peoplec=peoplec+1;
```

```
            if i=doselevelmtd then atmtd=atmtd+1;
            if i>doselevelmtd then abovemtd=abovemtd+1;
            if i<doselevelmtd then belowmtd=belowmtd+1;
            dltc=dltc+a[i,j];
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
      end;

      if sumi[i]=1 then do;
      do j=4 to 6;
            a[i,j]=rand('Bernoulli', dr);
            sumi[i]=sumi[i]+a[i,j];
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
          dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=atmtd+1;
            if i>doselevelmtd then abovemtd=abovemtd+1;
            if i<doselevelmtd then belowmtd=belowmtd+1;
            dltc=dltc+a[i,j];
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
      end;
      end;

end;
end;

dosel[k]=i;
dosemtd[k]=atmtd/peoplec*100;
doseover[k]=abovemtd/peoplec*100;
doseunder[k]=belowmtd/peoplec*100;
dltmtd[k]=dltatmtd/dltc*100;
dltover[k]=dltabovemtd/dltc*100;
dltunder[k]=dltbelowmtd/dltc*100;
dltoverall[k]=dltc/peoplec;

if atmtd ne 0 then mtddltrate[k]=dltatmtd/atmtd;
else mtddltrate[k]=0;
mtdpop[k]=atmtd;
totalpop[k]=peoplec;
totaldlt[k]=dltc;

end;

avgd=sumd/10000;
avgp=sump/10000;
ave= mean(of dosel1-dosel10000);
dev=STD(of dosel1-dosel10000);
serror=STDERR(of dosel1-dosel10000);
max= max(of dosel1-dosel10000);
```

```
median= median(of dosel1-dosel10000);
pctl_25=pctl(25, of dosel1-dosel10000);
pctl_50=pctl(50, of dosel1-dosel10000);
pctl_75=pctl(75, of dosel1-dosel10000);
avgdosemtd=mean(of dosemtd1-dosemtd10000);
avgdoseover=mean(of doseover1-doseover10000);
avgdoseunder=mean(of doseunder1-doseunder10000);
avgdltmtd=mean(of dltmtd1-dltmtd10000);
avgdltover=mean(of dltover1-dltover10000);
avgdltunder=mean(of dltunder1-dltunder10000);
avgdltrate=mean(of dltoverall1-dltoverall10000);
avgmtddltrate=mean(of mtddltrate1-mtddltrate10000);
avgmtdpop=median(of mtdpop1-mtdpop10000);
avgtotalpop=median(of totalpop1-totalpop10000);
avgtotalpop1=mean(of totalpop1-totalpop10000);
avgtotaldlt=median(of totaldlt1-totaldlt10000);
run;

/* print highest dose level reached/dose level at which the trial
stops, dose at highest dose level, dlt rate at the highest dose level,
factor to multiply present dose to get the next dose level, avg (+STD),
max and median number of dose levels explored, median and avg number of
patients on trial, median number of DLTs, avg DLT rate of the trial,
avg DLT rate at the MTD, median MTD pop, avg percentage of patients
dosed at MTD, under-dosed and over-dosed, avg number of patients and
DLTs at each dose level */

proc print;
var i dose dr frac ave dev max median avgtotalpop avgtotalpop1
avgtotaldlt avgdltrate avgmtddltrate avgmtdpop avgdosemtd avgdoseunder
avgdoseover peopledosel1-peopledosel10 dltdosel1-dltdosel10;
run;

data simi2;
set simi1;
array v_[*] dosel1-dosel10000;
array counts[10];
call missing (of counts[*]);
do i = 1 to dim(v_);
     counts[v_[i]] + 1;
end;
if counts1=. then counts1=0;
if counts2=. then counts2=0;
if counts3=. then counts3=0;
if counts4=. then counts4=0;
if counts5=. then counts5=0;
if counts6=. then counts6=0;
if counts7=. then counts7=0;
if counts8=. then counts8=0;
if counts9=. then counts9=0;
if counts10=. then counts10=0;

run;
```

```
/** number of times out of 10000 that the trial stops at each dose
level, from which the percentage of times each dose level is choosen as
the MTD can be determined - the MTD is one dose level below the dose
level at which the trial stops **/
proc print;
var counts1-counts10;
run;
```

SAS code for 3+3 and 2+4 designs where de-escalation is also allowed:

```
%macro alldesigns(design1, equation1);
data simi1;
call streaminit(1);

array a{1000,6};
array sumi{1000};
array sumitot{1000};
array dosel[10000];
array dosemtd{10000};
array doseover{10000};
array doseunder{10000};
array dltmtd{10000};
array dltover{10000};
array dltunder{10000};
array dltoverall{10000};
array mtddltrate{10000};
array mtdpop{10000};
array totalpop{10000};
array totaldlt{10000};
array peopledosel{10};
array dltdosel{10};
sumd=0;
sump=0;



/** program for either 3+3 or 2+4  design with deescalation **/

length design $10;
design="&design1";

length equation $20;
equation="&equation1";

if design='3+3' then do;
no1=3;
no2=4;
no3=6;
o1=   1;
o2=   3;
o3=   1;
```

```
o4=   6;
o5=   1;
o6=   6;
end;
if design='2+4' then do;
no1=2;
no2=3;
no3=6;
o1=   1;
o2=   2;
o3=   1;
o4=   6;
o5=   1;
o6=   6;
end;

startdose = 100;
mtddose=334;
dltrstartdose=0.01;
dltrmtd=0.2;
doselevelmtd=3;

/* program can handle one of 3 dose-toxicity curves */
if equation='linear' then do;

coeff2=1/(mtddose-startdose)*(dltrmtd-dltrstartdose);
coeff1= dltrstartdose - startdose*coeff2;

end;

if equation='logistic' then do;

coeff2=1/(mtddose-startdose)*( log(dltrmtd/(1-dltrmtd))-
log(dltrstartdose/(1-dltrstartdose)) );
coeff1= log(dltrstartdose/(1-dltrstartdose)) - startdose*coeff2;

end;

if equation='loglogistic' then do;

coeff2=1/(log(mtddose)-log(startdose))*( log(dltrmtd/(1-dltrmtd))-
log(dltrstartdose/(1-dltrstartdose)) );
coeff1= log(dltrstartdose/(1-dltrstartdose)) - log(startdose)*coeff2;

end;



do k=1 to 10000;


do i1=1 to 1000;
do j1=1 to 6;
a[i1,j1]=.;
```

```
sumi[i1]=.;
sumitot[i1]=.;
end;
end;

i=0;
gg=0;
sumi[1]=0;
sumitot[1]=0;
peoplec=0;
atmtd=0;
belowmtd=0;
abovemtd=0;
dltc=0;
dltatmtd=0;
dltbelowmtd=0;
dltabovemtd=0;

dose=100;

do until ((sumi[i]>o1 and sumins=o2) or (sumi[i]>o3 and sumins=o4));
    i = i+1;
      sumi[i]=0;
      sumitot[i]=0;
      sumins=0;


if i=1 then frac=1;
if i=2 then frac=2;
if i=3 then frac=1.67;
if i=4 then frac=1.5;
if i=5 then frac=1.4;
if i=6 then frac=1.33;
if i>=7 then frac=1.33;
dose=dose*frac;

if equation='linear' then dr=min(coeff1+coeff2*dose,1);

if equation='logistic' then dr=exp(coeff1+coeff2*dose)/(1+
exp(coeff1+coeff2*dose));

if equation='loglogistic' then dr=exp(coeff1+coeff2*log(dose))/(1+
exp(coeff1+coeff2*log(dose)));

      do j=1 to no1;
            a[i,j]=rand('Bernoulli', dr);
            sumi[i]=sumi[i]+a[i,j];
            sumitot[i]=sumitot[i]+1;
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            sumins=sumins+1;
            peoplec=peoplec+1;
```

```
            if i=doselevelmtd then atmtd=sum(atmtd,1);
            if i>doselevelmtd then abovemtd=sum(abovemtd,1);
            if i<doselevelmtd then belowmtd=sum(belowmtd,1);
            dltc=dltc+a[i,j];
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
      end;


if sumi[i]=1 then do;
do j=no2 to no3;
            a[i,j]=rand('Bernoulli', dr);
            sumi[i]=sumi[i]+a[i,j];
            sumitot[i]=sumitot[i]+1;
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            sumins=sumins+1;
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=sum(atmtd,1);
            if i>doselevelmtd then abovemtd=sum(abovemtd,1);
            if i<doselevelmtd then belowmtd=sum(belowmtd,1);
            dltc=dltc+a[i,j];
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
end;
end;


end;

***;
gg=i-1;

if gg>=1 then do;
/** stop if the dose level is 1 or (total number of patients at dose
level = 6 and total number of DLTs>1)or additional DLTs at dose level
is 0 or 1.**/
do until ( flaggg=1 or (sumitot[gg]=o4 and totdlt>o3) or totdlt in (0
1));
totdlt=0;
flaggg=.;
if gg=1 then flaggg=1;

if sumitot[gg]=o2 then do;


do jj=no2 to no3;
if equation='logistic' then do;
if gg=1 then dr1=0.01;
if gg=2 then dr1 =0.03828;
```

```
if gg=3 then dr1=0.2;
if gg=4 then dr1=0.71172;
if gg=5 then dr1=0.97471;
if gg=6 then dr1=0.99892;
if gg=7 then dr1=0.99998;
if gg>=8 then dr1=1;
end;
if equation='loglogistic' then do;
if gg=1 then dr1=0.01;
if gg=2 then dr1 =0.06004;
if gg=3 then dr1=0.2;
if gg=4 then dr1=0.42374;
if gg=5 then dr1=0.64287;
if gg=6 then dr1=0.79358;
if gg=7 then dr1=0.89143;
if gg=8 then dr1=0.94605;
if gg=9 then dr1=0.97399;
if gg=10 then dr1=0.98765;
end;
if equation='linear' then do;
if gg=1 then dr1=0.01;
if gg=2 then dr1 =0.0912;
if gg=3 then dr1=0.2;
if gg=4 then dr1=0.3356;
if gg=5 then dr1=0.49832;
if gg=6 then dr1=0.68626;
if gg=7 then dr1=0.93621;
if gg>=8 then dr1=1;
end;

            a[gg,jj]=rand('Bernoulli', dr1);
            sumi[gg]=sumi[gg]+a[gg,jj];
            sumitot[gg]=sumitot[gg]+1;
            totdlt=a[gg,jj]+totdlt;
            sumd=sumd+a[gg,jj];
            sump=sump+1;
            peopledosel[gg]=sum(peopledosel[gg],0.0001);
            dltdosel[gg]=sum(dltdosel[gg],a[gg,jj]/10000);
            peoplec=peoplec+1;
            if gg=doselevelmtd then atmtd=sum(atmtd,1);
            if gg>doselevelmtd then abovemtd=sum(abovemtd,1);
            if gg<doselevelmtd then belowmtd=sum(belowmtd,1);
            dltc=dltc+a[gg,jj];
            if gg=doselevelmtd then dltatmtd=dltatmtd+a[gg,jj];
            if gg>doselevelmtd then dltabovemtd=dltabovemtd+a[gg,jj];
            if gg<doselevelmtd then dltbelowmtd=dltbelowmtd+a[gg,jj];
end;

end;

if gg>1 then gg=gg-1;

end;
```

```
end;


***;
if gg=0 then dosel[k]=20;
else if gg=1 and flaggg=1 and totdlt in (0 1) then dosel[k]=gg;
else if gg=1 and flaggg=1 and totdlt in (2 3 4) then dosel[k]=20;
else if gg=1 and totdlt in (0 1) then dosel[k]=gg+1;
else if gg=1 and totdlt in (2 3 4) then dosel[k]=gg;
else if gg>1 and totdlt in (0 1) then dosel[k]=gg+1;
else if gg>1 and totdlt in (2 3 4) then dosel[k]=gg;

dosemtd[k]=atmtd/peoplec*100;
doseover[k]=abovemtd/peoplec*100;
doseunder[k]=belowmtd/peoplec*100;
dltmtd[k]=dltatmtd/dltc*100;
dltover[k]=dltabovemtd/dltc*100;
dltunder[k]=dltbelowmtd/dltc*100;

dltoverall[k]=dltc/peoplec;
if atmtd ne 0 then mtddltrate[k]=dltatmtd/atmtd;
else mtddltrate[k]=0;
mtdpop[k]=atmtd;
totalpop[k]=peoplec;
totaldlt[k]=dltc;

end;
avgd=sumd/10000;
avgp=sump/10000;
ave= mean(of dosel1-dosel10000);
dev=STD(of dosel1-dosel10000);
serror=STDERR(of dosel1-dosel10000);
max= max(of dosel1-dosel10000);
median= median(of dosel1-dosel10000);
pctl_25=pctl(25, of dosel1-dosel10000);
pctl_50=pctl(50, of dosel1-dosel10000);
pctl_75=pctl(75, of dosel1-dosel10000);
avgdosemtd=mean(of dosemtd1-dosemtd10000);
avgdoseover=mean(of doseover1-doseover10000);
avgdoseunder=mean(of doseunder1-doseunder10000);
avgdltmtd=mean(of dltmtd1-dltmtd10000);
avgdltover=mean(of dltover1-dltover10000);
avgdltunder=mean(of dltunder1-dltunder10000);
avgdltrate=mean(of dltoverall1-dltoverall10000);
avgmtddltrate=mean(of mtddltrate1-mtddltrate10000);
avgmtdpop=median(of mtdpop1-mtdpop10000);
avgtotalpop=median(of totalpop1-totalpop10000);
avgtotalpop1=mean(of totalpop1-totalpop10000);
avgtotaldlt=median(of totaldlt1-totaldlt10000);
run;

/* print highest dose level reached, dose level at which the simulated
trial finally stopped, total number of additional DLTs in the final
dose level at which the trial stopped, median and avg number of
```

```
patients on trial, median number of DLTs, avg DLT rate of the trial,
avg DLT rate at the MTD, avg percentage of patients dosed at MTD,
under-dosed and over-dosed, avg number of patients and DLTs at each
dose level */

proc print;
var i gg totdlt avgtotalpop avgtotalpop1 avgtotaldlt avgdltrate
avgmtddltrate avgdosemtd avgdoseunder avgdoseover peopledosel1-
peopledosel10 dltdosel1-dltdosel10;
run;

data simi2;
set simi1;
array v_[*] dosel1-dosel10000;
array counts[20];
call missing (of counts[*]);
do i = 1 to dim(v_);
     counts[v_[i]] + 1;
end;
if counts1=. then counts1=0;
if counts2=. then counts2=0;
if counts3=. then counts3=0;
if counts4=. then counts4=0;
if counts5=. then counts5=0;
if counts6=. then counts6=0;
if counts7=. then counts7=0;
if counts8=. then counts8=0;
if counts9=. then counts9=0;
if counts10=. then counts10=0;

run;

/* times out of 10000 that each dose level is choosen as the MTD */

proc print;
var counts1-counts20;
run;

%mend alldesigns;

%alldesigns(3+3, logistic);
```

SAS code for starting at lower dose levels below the MTD as well as code for the linear dose-toxicity curves starting with different offsets:

```
%macro alldesigns(design1, equation1);
data simi1;
call streaminit(1);

array a{1000,6};
array sumi{1000};
array dosel[10000];
array dosemtd{10000};
```

```
array doseover{10000};
array doseunder{10000};
array dltmtd{10000};
array dltover{10000};
array dltunder{10000};
array dltoverall{10000};
array mtddltrate{10000};
array mtdpop{10000};
array totalpop{10000};
array totaldlt{10000};
array peopledosel{20};
array dltdosel{20};
sumd=0;
sump=0;


/** design can be  3+3, 2+4, 4+4 a,4+4 b, 5+5 a, 5+5 b,5+5 c, 3+3+3,
3+1+1, 10+10 a, or 20+20 a**/

length design $10;
design="&design1";

length equation $20;
equation="&equation1";

if design='3+3' then do;
no1=3;
no2=4;
no3=6;
o1=   1;
o2=   3;
o3=   1;
o4=   6;
o5=   1;
o6=   6;
end;
if design='2+4' then do;
no1=2;
no2=3;
no3=6;
o1=   1;
o2=   2;
o3=   1;
o4=   6;
o5=   1;
o6=   6;
end;
if design='4+4a' then do;
no1=4;
no2=5;
no3=8;
o1=   2;
o2=   4;
o3=   2;
```

```
o4=   8;
o5=   2;
o6=   8;
end;
if design='4+4b' then do;
no1=4;
no2=5;
no3=8;
o1=   2;
o2=   4;
o3=   3;
o4=   8;
o5=   3;
o6=   8;
end;

if design='3+1+1' then do;
no1=3;
o1=   2;
o2=   3;
o3=   2;
o4=   4;
o5=   2;
o6=   5;
end;
if design='3+3+3' then do;
no1=3;
no2=4;
no3=6;
o1=   1;
o2=   3;
o3=   2;
o4=   6;
o5=   2;
o6=   9;
end;
if design='5+5a' then do;
no1=5;
no2=6;
no3=10;
o1=   1;
o2=   5;
o3=   1;
o4=   10;
o5=   1;
o6=   10;
end;
if design='5+5b' then do;
no1=5;
no2=6;
no3=10;
o1=   2;
o2=   5;
o3=   2;
```

```
o4=   10;
o5=   2;
o6=   10;
end;
if design='5+5c' then do;
no1=5;
no2=6;
no3=10;
o1=   2;
o2=   5;
o3=   3;
o4=   10;
o5=   3;
o6=   10;
end;
if design='10+10a' then do;
no1=10;
no2=11;
no3=20;
o1=   4;
o2=   10;
o3=   4;
o4=   20;
o5=   4;
o6=   20;
end;
if design='20+20a' then do;
no1=20;
no2=21;
no3=40;
o1=   8;
o2=   20;
o3=   8;
o4=   40;
o5=   8;
o6=   40;
end;

startdose = 100;
mtddose=334;
dltrstartdose=0.01;
dltrmtd=0.2;

/* program can handle one of 3 dose-toxicity curves */
if equation='linear' then do;

coeff2=1/(mtddose-startdose)*(dltrmtd-dltrstartdose);
coeff1= dltrstartdose - startdose*coeff2;

end;

if equation='logistic' then do;
```

```
coeff2=1/(mtddose-startdose)*( log(dltrmtd/(1-dltrmtd))-
log(dltrstartdose/(1-dltrstartdose)) );
coeff1= log(dltrstartdose/(1-dltrstartdose)) - startdose*coeff2;

end;

if equation='loglogistic' then do;

coeff2=1/(log(mtddose)-log(startdose))*( log(dltrmtd/(1-dltrmtd))-
log(dltrstartdose/(1-dltrstartdose)) );
coeff1= log(dltrstartdose/(1-dltrstartdose)) - log(startdose)*coeff2;

end;



do k=1 to 10000;


do i1=1 to 1000;
do j1=1 to 6;
a[i1,j1]=.;
sumi[i1]=.;
end;
end;

i=0;
sumi[1]=0;
peoplec=0;
atmtd=0;
belowmtd=0;
abovemtd=0;
dltc=0;
dltatmtd=0;
dltbelowmtd=0;
dltabovemtd=0;

dose=100;
/** ii2 helps define how many dose levels above the lowest dose level
the true MTD is. For our simulation study, ii2 can be 3, 2, 1 or 0 **/
ii2=3;
doselevelmtd=ii2+3;

do until ((sumi[i]>o1 and sumins=o2) or (sumi[i]>o3 and sumins=o4) or
(sumi[i]>o5 and sumins=o6));
    i = i+1;
      sumi[i]=0;
      sumins=0;

/** You need one of the below 3 sets of numbers for the 3 linear dose
toxicity curves with different offsets considered to examine the effect
of the location of the starting dose relative to the true MTD.
Otherwise comment these 3 sets of numbers. The "i"s below are such that
we start 6 doses levels below the true MTD in each case, and will need
```

```
to be changed if we start 5 dose levels below the true MTD for example.
**/

/*
if i<=6 then dr=0.1;
if i=7 then dr=0.2;
if i=8 then dr=0.32;
if i=9 then dr=0.4585988;
if i=10 then dr=0.6429401;
if i=11 then dr=0.8881078;
if i>=12 then dr=1;
*/


if i<=5 then dr=0.05;
if i=6 then dr=0.1181818;
if i=7 then dr=0.2;
if i=8 then dr=0.2944992;
if i=9 then dr=0.4201865;
if i=10 then dr=0.5873463;
if i=11 then dr=0.8096721;
if i>=12 then dr=1;


/*
if i<=4 then dr=0.0;
if i=5 then dr=0.05577262;
if i=6 then dr=0.1226998;
if i=7 then dr=0.2;
if i=8 then dr=0.302812;
if i=9 then dr=0.4395485;
if i=10 then dr=0.6214107;
if i=11 then dr=0.8632836;
if i>=12 then dr=1;
*/

/** If you are examining the effect of the location of the starting
dose relative to the true MTD for the logistic or log-logistic dose-
toxicitiy curves, uncomment the below section **/

/*

if i=ii2-2 then dose=12.5;
if i=ii2-1 then dose=25;
if i=ii2 then dose=50;
if ii2=0 and i=ii2+1 then frac=1;
if ii2>0 and i=ii2+1 then frac=2;
if i=ii2+2 then frac=2;
if i=ii2+3 then frac=1.67;
if i=ii2+4 then frac=1.5;
if i=ii2+5 then frac=1.4;
if i>=ii2+6 then frac=1.33;
if i>=ii2+1 then dose=dose*frac;
```

```
if equation='linear' then dr=min(coeff1+coeff2*dose,1);

if equation='logistic' then dr=exp(coeff1+coeff2*dose)/(1+
exp(coeff1+coeff2*dose));

if equation='loglogistic' then dr=exp(coeff1+coeff2*log(dose))/(1+
exp(coeff1+coeff2*log(dose)));

*/

      do j=1 to no1;
            a[i,j]=rand('Bernoulli', dr);
            sumi[i]=sumi[i]+a[i,j];
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            sumins=sumins+1;
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=sum(atmtd,1);
            if i>doselevelmtd then abovemtd=sum(abovemtd,1);
            if i<doselevelmtd then belowmtd=sum(belowmtd,1);
            dltc=dltc+a[i,j];
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
      end;

if (design='3+3' or design='2+4' or design='3+3+3' or design='5+5a')
then do;
if sumi[i]=1 then do;
do j=no2 to no3;
            a[i,j]=rand('Bernoulli', dr);
            sumi[i]=sumi[i]+a[i,j];
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            sumins=sumins+1;
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=sum(atmtd,1);
            if i>doselevelmtd then abovemtd=sum(abovemtd,1);
            if i<doselevelmtd then belowmtd=sum(belowmtd,1);
            dltc=dltc+a[i,j];
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
end;
end;
end;

if (design='4+4a' or design='5+5b') then do;
if (sumi[i]=1 or sumi[i]=2) then do;
do j=no2 to no3;
```

```
            a[i,j]=rand('Bernoulli', dr);
            sumi[i]=sumi[i]+a[i,j];
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            sumins=sumins+1;
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=sum(atmtd,1);
            if i>doselevelmtd then abovemtd=sum(abovemtd,1);
            if i<doselevelmtd then belowmtd=sum(belowmtd,1);
            dltc=dltc+a[i,j];
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
end;
end;
end;

if (design='4+4b' or design='5+5c') then do;
if sumi[i]=2 then do;
do j=no2 to no3;
            a[i,j]=rand('Bernoulli', dr);
            sumi[i]=sumi[i]+a[i,j];
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            sumins=sumins+1;
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=sum(atmtd,1);
            if i>doselevelmtd then abovemtd=sum(abovemtd,1);
            if i<doselevelmtd then belowmtd=sum(belowmtd,1);
            dltc=dltc+a[i,j];
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
end;
end;
end;

if (design='10+10a') then do;
if (sumi[i]=3 or sumi[i]=4) then do;
do j=no2 to no3;
            a[i,j]=rand('Bernoulli', dr);
            sumi[i]=sumi[i]+a[i,j];
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            sumins=sumins+1;
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=sum(atmtd,1);
            if i>doselevelmtd then abovemtd=sum(abovemtd,1);
```

```
            if i<doselevelmtd then belowmtd=sum(belowmtd,1);
            dltc=dltc+a[i,j];
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
end;
end;
end;

if (design='20+20a') then do;
if (sumi[i]=7 or sumi[i]=8) then do;
do j=no2 to no3;
            a[i,j]=rand('Bernoulli', dr);
            sumi[i]=sumi[i]+a[i,j];
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            sumins=sumins+1;
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=sum(atmtd,1);
            if i>doselevelmtd then abovemtd=sum(abovemtd,1);
            if i<doselevelmtd then belowmtd=sum(belowmtd,1);
            dltc=dltc+a[i,j];
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
end;
end;
end;

if design='3+3+3' then do;
if (sumi[i]=2 and sumins=6) then do;
do j=7 to 9;
            a[i,j]=rand('Bernoulli', dr);
            sumi[i]=sumi[i]+a[i,j];
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            sumins=sumins+1;
            peoplec=peoplec+1;
      if i=doselevelmtd then atmtd=sum(atmtd,1);
            if i>doselevelmtd then abovemtd=sum(abovemtd,1);
            if i<doselevelmtd then belowmtd=sum(belowmtd,1);
            dltc=dltc+a[i,j];
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
end;
end;
end;

if design='3+1+1' then do;
```

```
if ((sumi[i]=1) or (sumi[i]=2)) then do;
            a[i,4]=rand('Bernoulli', dr);
            sumi[i]=sumi[i]+a[i,4];
            sumd=sumd+a[i,4];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            sumins=sumins+1;
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=sum(atmtd,1);
            if i>doselevelmtd then abovemtd=sum(abovemtd,1);
            if i<doselevelmtd then belowmtd=sum(belowmtd,1);
            dltc=dltc+a[i,j];
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
end;

if ((sumi[i]=2) and (sumins=4)) then do;
            a[i,5]=rand('Bernoulli', dr);
            sumi[i]=sumi[i]+a[i,5];
            sumd=sumd+a[i,5];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            sumins=sumins+1;
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=sum(atmtd,1);
            if i>doselevelmtd then abovemtd=sum(abovemtd,1);
            if i<doselevelmtd then belowmtd=sum(belowmtd,1);
            dltc=dltc+a[i,j];
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
end;
end;


end;
dosel[k]=i;
dosemtd[k]=atmtd/peoplec*100;
doseover[k]=abovemtd/peoplec*100;
doseunder[k]=belowmtd/peoplec*100;
dltmtd[k]=dltatmtd/dltc*100;
dltover[k]=dltabovemtd/dltc*100;
dltunder[k]=dltbelowmtd/dltc*100;

dltoverall[k]=dltc/peoplec;
if atmtd ne 0 then mtddltrate[k]=dltatmtd/atmtd;
else mtddltrate[k]=0;
mtdpop[k]=atmtd;
totalpop[k]=peoplec;
totaldlt[k]=dltc;
```

```
end;
avgd=sumd/10000;
avgp=sump/10000;
ave= mean(of dosel1-dosel10000);
dev=STD(of dosel1-dosel10000);
serror=STDERR(of dosel1-dosel10000);
max= max(of dosel1-dosel10000);
median= median(of dosel1-dosel10000);
pctl_25=pctl(25, of dosel1-dosel10000);
pctl_50=pctl(50, of dosel1-dosel10000);
pctl_75=pctl(75, of dosel1-dosel10000);
avgdosemtd=mean(of dosemtd1-dosemtd10000);
avgdoseover=mean(of doseover1-doseover10000);
avgdoseunder=mean(of doseunder1-doseunder10000);
avgdltmtd=mean(of dltmtd1-dltmtd10000);
avgdltover=mean(of dltover1-dltover10000);
avgdltunder=mean(of dltunder1-dltunder10000);
avgdltrate=mean(of dltoverall1-dltoverall10000);
avgmtddltrate=mean(of mtddltrate1-mtddltrate10000);
avgmtdpop=median(of mtdpop1-mtdpop10000);
avgtotalpop=median(of totalpop1-totalpop10000);
avgtotalpop1=mean(of totalpop1-totalpop10000);
avgtotaldlt=median(of totaldlt1-totaldlt10000);
run;

/* print design, highest dose level reached/dose level at which the
trial stops, dlt rate at the highest dose level, avg (+STD), max and
median number of dose levels explored, median and avg number of
patients on trial, median number of DLTs, avg DLT rate of the trial,
avg DLT rate at the MTD, median MTD pop, avg percentage of patients
dosed at MTD, under-dosed and over-dosed, avg number of patients and
DLTs at each dose level */

proc print;
var design i dr ave dev max median avgtotalpop avgtotalpop1 avgtotaldlt
avgdltrate avgmtddltrate avgmtdpop avgdosemtd avgdoseunder avgdoseover
peopledosel1-peopledosel10 dltdosel1-dltdosel10;
run;

data simi2;
set simi1;
array v_[*] dosel1-dosel10000;
array counts[20];
call missing (of counts[*]);
do i = 1 to dim(v_);
     counts[v_[i]] + 1;
end;

run;

/* times out of 10000 that the trial stops at each dose level, from
which the percentage of times each dose level is choosen as the MTD can
be determined - the MTD is one dose level below the dose level at which
the trial stops */
```

```
proc print;
var counts1-counts20;
run;

%mend alldesigns;

%alldesigns(3+3, logistic);
```

R code for TEQR design using the R package TEQR:

```
# logistic dose-toxicity curve
set.seed(1111)
OperChar<-teqrOCtox(
sim=2000,
firstdose=1,
probt=c(.01,.03828,.2,.71172,.97471),
cohortSize=3,
MaxNoCohorts=30,
MTDss=12,
pTarget=.2,
eq1=.05,
eq2=.05,
tootoxic=.34)
OperChar
```

R code for BOIN design using the R package BOIN:

```
# logistic dose-toxicity curve
```
get.oc(target=0.2, p.true=c(0.01, 0.03828, 0.2, 0.71172, 0.97471),startdose=1, ncohort=7, cohortsize=3, ntrial=2000)
get.boundary(target=0.2, ncohort=7, cohortsize=3)

For the mTPI design,we used the R program given in the following weblink.
http://health.bsd.uchicago.edu/yji/software2.htm

For the CRM design, we used the CRM Simulator software.
https://biostatistics.mdanderson.org/SoftwareDownload/SingleSoftware.aspx?Software_Id=13

For the EWOC design, we used the Web-EWOC software.
https://biostatistics.csmc.edu/ewoc/ewocWeb.php

/* SAS Code for finding the probability of being the highest dose level examined for the 3+3 design and some of its extensions (escalation-only designs) */

```
data a2;
a=1;
```

```
b=1;
c=1;
d=1;
e=1;
f=1;
g=1;
h=1;
do k=1 to 6;
p=0.05+(k-1)*0.05;
q=1-p;
/* 3+3 */
vara = (q**3) + (3*p*(q**5));
varb = 1-vara;
varc = 3*q*p**2+p**3 + 9*p**2*q**4 + 9*p**3*q**3 +3*q**2*p**4;
a=a*vara;
fin1=a*varb/vara;
/* 3+3+3 */
vard=q**3+3*p*q**5+9*p**2*q**7;
vare=1-vard;
varf = 3*q*p**2+p**3 +3*q**2*p**4 + 9*p**3*q**3+ 27*p**3*q**6 +
27*p**4*q**5 + 9*p**5*q**4;
b=b*vard;
fin2=b*vare/vard;
/* 2+4 */
varg = (q**2) + (2*p*(q**5));
varh = 1-varg;
vari = p**2 + 8*p**2*q**4+12*p**3*q**3+8*p**4*q**2+2*p**5*q;
c=c*varg;
fin3=c*varh/varg;
/* 4+4 a */
varj=q**4+4*p*q**7+22*p**2*q**6;
vark=1-varj;
varl=48*p**3*q**5+52*p**4*q**4+28*p**5*q**3+6*p**6*q**2+4*p**3*q+p**4;
d=d*varj;
fin4=d*vark/varj;
/* 4+4 b*/
varm = (q**4) + (4*p*(q**3)) + 6*p**2*q**6 + 24*p**3*q**5;
varn = 1-varm;
varo = 4*p**3*q+p**4+36*p**4*q**4+24*p**5*q**3+6*p**6*q**2;
e=e*varm;
fin5=e*varn/varm;
/* 5+5 a */
varp = (q**5) + (5*p*(q**9)) + 35*p**2*q**8;
varq = 1-varp;
varr =
p**5+5*p**4*q+10*p**3*q**2+100*p**3*q**7+150*p**4*q**6+125*p**5*q**5+55
*p**6*q**4+10*p**7*q**3;
f=f*varp;
fin6=f*varq/varp;
/* 5+5 b */
vars = (q**5) + (5*p*(q**4)) + 10*p**2*q**8 + 50*p**3*q**7;
vart = 1-vars;
```

```
varu =
p**5+5*p**4*q+10*p**3*q**2+100*p**4*q**6+100*p**5*q**5+50*p**6*q**4+10*
p**7*q**3;
g=g*vars;
fin7=g*vart/vars;
/* 3+1+1 */
varv = q**3+3*p*q**3+6*p**2*q**3;
varw = 1-varv;
varx=p**3+3*p**3*q+6*p**3*q**2;
h=h*varv;
fin8=h*varw/varv;
output;
end;

proc print;
var p fin1-fin8;
title 'Probabilities of being the highest dose level examined for
various designs that are extensions of the 3+3 design and that consider
only escalation';
run
```

Codes for Chapter 3

SAS code for 20+20 accelerated titration design where there is a stopping rule after every 6 or 8 patients are dosed:

```
/** The upper limit for toxicity in the Bayesian decision rule for safety
is 0.33 and the lower limit for efficacy in the Bayesian decision rule
for efficacy is 0.5. For other thresholds, these numbers need to be
changed in the code below. **/

%macro alldesigns(design1, equation1);
data simi1;
call streaminit(1);

array a{1000,40};
array aeff{1000,40};
array asucc{1000,40};
array sumi{1000};
array sumieff{1000};
array sumins{1000};
array flag{1000};
array flagabc{1000};
array flagabcd{1000};
array flagabcdee{1000};
array flagabcdff{1000};
array flagabcdgg{1000};
array flagabcdhh{1000};
array bayflag{1000};
```

```
array pvaluea{1000};
array pvalueb{1000};

array dosel{10000};
array dosemtd{10000};
array doseover{10000};
array doseunder{10000};
/*
array dltmtd{11};
array dltover{11};
array dltunder{11};
*/
array dltoverall{10000};
array mtddltrate{10000};
array mtdpop{10000};
array totalpop{10000};
array totaldlt{10000};
array index{10000};
array peopledosel{10};
array dltdosel{10};
array onlyrespsdosel{10};
array respsnodltsdosel{10};
array dltratedoselevel{10};
array responseratedoselevel{10};
array utilitydltrate{10};
array utilityresponserate{10};
array utility{10};
array ratio{10};
array ratiof{10};

sumd=0;
sump=0;
count9=0;

length design $10;
design="&design1";

length equation $20;
equation="&equation1";

if design='20+20' then do;
no1=20;
no2=21;
no3=40;
no4=14;
no5=15;
o1=   8;
o2=   20;
o3=   8;
o4=   40;
end;

/** parameters used to define the dose-toxicity curve **/
startdose = 100;
```

```
mtddose=501;
dltrstartdose=0.01;
dltrmtd=0.2;
doselevelmtd=4;


if equation='linear' then do;

coeff2=1/(mtddose-startdose)*(dltrmtd-dltrstartdose);
coeff1= dltrstartdose - startdose*coeff2;

end;

if equation='logistic' then do;

coeff2=1/(mtddose-startdose)*( log(dltrmtd/(1-dltrmtd))-
log(dltrstartdose/(1-dltrstartdose)) );
coeff1= log(dltrstartdose/(1-dltrstartdose)) - startdose*coeff2;

end;

if equation='loglogistic' then do;

coeff2=1/(log(mtddose)-log(startdose))*( log(dltrmtd/(1-dltrmtd))-
log(dltrstartdose/(1-dltrstartdose)) );
coeff1= log(dltrstartdose/(1-dltrstartdose)) - log(startdose)*coeff2;

end;


do k=1 to 10000;


do i1=1 to 1000;
do j1=1 to 40;
a[i1,j1]=.;
aeff[i1,j1]=.;
asucc[i1,j1]=.;
sumi[i1]=.;
sumieff[i1]=.;
sumins[i1]=.;
flag[i1]=.;
flagabc[i1]=.;
flagabcd[i1]=.;
flagabcdee[i1]=.;
flagabcdff[i1]=.;
flagabcdgg[i1]=.;
flagabcdgg[i1]=.;

pvaluea[i1]=.;
pvalueb[i1]=.;

end;
end;
```

```
i=0;
sumi[1]=0;
sumieff[1]=0;
sumins[1]=0;
peoplec=0;
atmtd=0;
belowmtd=0;
abovemtd=0;
dltc=0;
dltatmtd=0;
dltbelowmtd=0;
dltabovemtd=0;


dose=100;

do until (sumi[i]>=1);
    i = i+1;
      sumi[i]=0;
      sumieff[i]=0;
      sumins[i]=0;

    if i=1 then frac=1;
    if i=2 then frac=2;
    if i=3 then frac=1.67;
    if i=4 then frac=1.5;
    if i=5 then frac=1.4;
    if i=6 then frac=1.33;
    if i>=7 then frac=1.33;

    dose=dose*frac;

    if equation='linear' then dr=min(coeff1+coeff2*dose,1);

    if equation='logistic' then dr=exp(coeff1+coeff2*dose)/(1+
exp(coeff1+coeff2*dose));

    if equation='loglogistic' then dr=exp(coeff1+coeff2*log(dose))/(1+
exp(coeff1+coeff2*log(dose)));

/** true underlying response rates **/
    if i=1 then dreff=0.01;
    if i=2 then dreff=0.05;
    if i=3 then dreff=0.15;
    if i=4 then dreff=0.45;
    if i=5 then dreff=0.2;
    if i>=6 then dreff=0.05;

/** correlation coefficient - correlation between true toxicity and
efficacy rates **/
      r=0;

    p1=dr;
    p2=dreff;
```

```
/** start with 3 patients at each dose level until at least 1 DLT is
observed **/
      do j=1 to 3;
      a[i,j]=rand('Bernoulli', dr);
      q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
      q2=(p2-q1*p1)/(1-p1);
    if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
    else aeff[i,j]=rand("binomial",q2,1);
    if a[i,j]=0 and aeff[i,j]=1 then asucc[i,j]=1;
    else asucc[i,j]=0;
      sumi[i]=sumi[i]+a[i,j];
      sumieff[i]=sumieff[i]+aeff[i,j];
      sumd=sumd+a[i,j];
      sump=sump+1;
      peopledosel[i]=sum(peopledosel[i],0.0001);
      dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
      onlyrespsdosel[i]=sum(onlyrespsdosel[i],aeff[i,j]/10000);
      respsnodltsdosel[i]=sum(respsnodltsdosel[i],asucc[i,j]/10000);
      sumins[i]=sumins[i]+1;
    peoplec=peoplec+1;
      if i=doselevelmtd then atmtd=atmtd+1;
      if i>doselevelmtd then abovemtd=abovemtd+1;
      if i<doselevelmtd then belowmtd=belowmtd+1;
      dltc=dltc+a[i,j];
      /*
      if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
      if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
      if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
      */
      pvaluea[i]=cdf("beta", 0.33, 0.5+sumi[i], 0.5+sumins[i]-sumi[i]);
      pvalueb[i]=1-cdf("beta", 0.5, 0.5+sumieff[i], 0.5+sumins[i]-
sumieff[i]);
      end;

end;

/** switch to cohorts of 20 in groups of 6 or 8 patients once one or
more DLTs are observed among 3 patients **/
do j=4 to 6;
       a[i,j]=rand('Bernoulli', dr);
       q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
       q2=(p2-q1*p1)/(1-p1);
       if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
       else aeff[i,j]=rand("binomial",q2,1);
           if a[i,j]=0 and aeff[i,j]=1 then asucc[i,j]=1;
           else asucc[i,j]=0;
           sumi[i]=sumi[i]+a[i,j];
           sumieff[i]=sumieff[i]+aeff[i,j];
           sumd=sumd+a[i,j];
           sump=sump+1;
           peopledosel[i]=sum(peopledosel[i],0.0001);
           dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
           onlyrespsdosel[i]=sum(onlyrespsdosel[i],aeff[i,j]/10000);
```

```
respsnodltsdosel[i]=sum(respsnodltsdosel[i],asucc[i,j]/10000);
            sumins[i]=sumins[i]+1;
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=atmtd+1;
            if i>doselevelmtd then abovemtd=abovemtd+1;
            if i<doselevelmtd then belowmtd=belowmtd+1;
            dltc=dltc+a[i,j];
            /*
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
            */
            pvaluea[i]=cdf("beta", 0.33, 0.5+sumi[i], 0.5+sumins[i]-
sumi[i]);
            pvalueb[i]=1-cdf("beta", 0.5, 0.5+sumieff[i],
0.5+sumins[i]-sumieff[i]);
end;

if sumi[i]<4 then flagabcd[i]=1;
else flagabcd[i]=0;
jj3=i;

/** proceed to add 8 more patients at the same dose level if there are
less than 4 DLTs in the first 6 patients in the dose level **/

if flagabcd[jj3]=1 then do;

do j=7 to 14;
            a[i,j]=rand('Bernoulli', dr);
            q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
          q2=(p2-q1*p1)/(1-p1);
        if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
        else aeff[i,j]=rand("binomial",q2,1);
            if a[i,j]=0 and aeff[i,j]=1 then asucc[i,j]=1;
            else asucc[i,j]=0;
            sumi[i]=sumi[i]+a[i,j];
            sumieff[i]=sumieff[i]+aeff[i,j];
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            onlyrespsdosel[i]=sum(onlyrespsdosel[i],aeff[i,j]/10000);

respsnodltsdosel[i]=sum(respsnodltsdosel[i],asucc[i,j]/10000);
            sumins[i]=sumins[i]+1;
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=atmtd+1;
            if i>doselevelmtd then abovemtd=abovemtd+1;
            if i<doselevelmtd then belowmtd=belowmtd+1;
            dltc=dltc+a[i,j];
            /*
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
```

```
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
            */
            pvaluea[i]=cdf("beta", 0.33, 0.5+sumi[i], 0.5+sumins[i]-
sumi[i]);
            pvalueb[i]=1-cdf("beta", 0.5, 0.5+sumieff[i],
0.5+sumins[i]-sumieff[i]);
end;

/** proceed to add 6 more patients at the same dose level if there is
at least one response in the first 14 patients and the number of DLTs
is <9 or there is at least 1 DLT but less than 9 DLTs  **/

if ((0<sumi[i]<9 or (sumieff[i]>0 and 0<=sumi[i]<9))) then do;
do j=15 to 20;

            a[i,j]=rand("binomial",dr,1);
        q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
          q2=(p2-q1*p1)/(1-p1);
        if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
        else aeff[i,j]=rand("binomial",q2,1);
            if a[i,j]=0 and aeff[i,j]=1 then asucc[i,j]=1;
            else asucc[i,j]=0;
            sumi[i]=sumi[i]+a[i,j];
            sumieff[i]=sumieff[i]+aeff[i,j];
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            onlyrespsdosel[i]=sum(onlyrespsdosel[i],aeff[i,j]/10000);

        respsnodltsdosel[i]=sum(respsnodltsdosel[i],asucc[i,j]/10000);
            sumins[i]=sumins[i]+1;
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=sum(atmtd,1);
            if i>doselevelmtd then abovemtd=sum(abovemtd,1);
            if i<doselevelmtd then belowmtd=sum(belowmtd,1);
            dltc=dltc+a[i,j];
/*
        if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
*/
            pvaluea[i]=cdf("beta", 0.33, 0.5+sumi[i], 0.5+sumins[i]-
sumi[i]);
          pvalueb[i]=1-cdf("beta", 0.5, 0.5+sumieff[i], 0.5+sumins[i]-
sumieff[i]);
 end;
 end;

/** if number of DLTs in the first 20 patients is 7 or 8 then add
patients in cohorts of 6 or 8 for a max of 40 patients, as long as less
than 9 DLTs are observed **/
if sumi[i]=7 or sumi[i]=8 then do;
```

```
do j=21 to 26;
            a[i,j]=rand('Bernoulli', dr);
            q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
          q2=(p2-q1*p1)/(1-p1);
        if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
        else aeff[i,j]=rand("binomial",q2,1);
            if a[i,j]=0 and aeff[i,j]=1 then asucc[i,j]=1;
            else asucc[i,j]=0;
            sumi[i]=sumi[i]+a[i,j];
            sumieff[i]=sumieff[i]+aeff[i,j];
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            onlyrespsdosel[i]=sum(onlyrespsdosel[i],aeff[i,j]/10000);

      respsnodltsdosel[i]=sum(respsnodltsdosel[i],asucc[i,j]/10000);
            sumins[i]=sumins[i]+1;
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=atmtd+1;
            if i>doselevelmtd then abovemtd=abovemtd+1;
            if i<doselevelmtd then belowmtd=belowmtd+1;
            dltc=dltc+a[i,j];
            /*
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
            */
            pvaluea[i]=cdf("beta", 0.33, 0.5+sumi[i], 0.5+sumins[i]-
sumi[i]);
          pvalueb[i]=1-cdf("beta", 0.5, 0.5+sumieff[i], 0.5+sumins[i]-
sumieff[i]);
end;

if sumi[i]<9 then flagabcdee[i]=1;
else flagabcdee[i]=0;
jj26=i;

if flagabcdee[jj26]=1 then do;
do j=27 to 34;
            a[i,j]=rand('Bernoulli', dr);
            q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
          q2=(p2-q1*p1)/(1-p1);
        if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
        else aeff[i,j]=rand("binomial",q2,1);
            if a[i,j]=0 and aeff[i,j]=1 then asucc[i,j]=1;
            else asucc[i,j]=0;
            sumi[i]=sumi[i]+a[i,j];
            sumieff[i]=sumieff[i]+aeff[i,j];
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            onlyrespsdosel[i]=sum(onlyrespsdosel[i],aeff[i,j]/10000);
```

```
      respsnodltsdosel[i]=sum(respsnodltsdosel[i],asucc[i,j]/10000);
            sumins[i]=sumins[i]+1;
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=atmtd+1;
            if i>doselevelmtd then abovemtd=abovemtd+1;
            if i<doselevelmtd then belowmtd=belowmtd+1;
            dltc=dltc+a[i,j];
            /*
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
            */
            pvaluea[i]=cdf("beta", 0.33, 0.5+sumi[i], 0.5+sumins[i]-
sumi[i]);
          pvalueb[i]=1-cdf("beta", 0.5, 0.5+sumieff[i], 0.5+sumins[i]-
sumieff[i]);
end;

if sumi[i]<9 then flagabcdff[i]=1;
else flagabcdff[i]=0;
jj35=i;

if flagabcdff[jj35]=1 then do;
do j=35 to 40;
            a[i,j]=rand('Bernoulli', dr);
            q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
          q2=(p2-q1*p1)/(1-p1);
        if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
        else aeff[i,j]=rand("binomial",q2,1);
            if a[i,j]=0 and aeff[i,j]=1 then asucc[i,j]=1;
            else asucc[i,j]=0;
            sumi[i]=sumi[i]+a[i,j];
            sumieff[i]=sumieff[i]+aeff[i,j];
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            onlyrespsdosel[i]=sum(onlyrespsdosel[i],aeff[i,j]/10000);

      respsnodltsdosel[i]=sum(respsnodltsdosel[i],asucc[i,j]/10000);
            sumins[i]=sumins[i]+1;
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=atmtd+1;
            if i>doselevelmtd then abovemtd=abovemtd+1;
            if i<doselevelmtd then belowmtd=belowmtd+1;
            dltc=dltc+a[i,j];
            /*
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
            */
            pvaluea[i]=cdf("beta", 0.33, 0.5+sumi[i], 0.5+sumins[i]-
sumi[i]);
```

```
          pvalueb[i]=1-cdf("beta", 0.5, 0.5+sumieff[i], 0.5+sumins[i]-
sumieff[i]);
end;
end;

end;

end;

end;

if sumi[i]<=8 then flag[i]=1;
else flag[i]=0;
jj1=i;

/** one can escalate if there are <=6 DLTs in the first 20 patients in
a dose level, <=8 DLTs in 40 patients at a dose level or if there are
no DLTs and no responses in the first 14 patients in a dose level **/
if flag[jj1]=1 and flagabcd[jj3]=1 then do;

do until (sumi[i]>o1 or flagabc[jj2]=0);
    i = i+1;
      sumi[i]=0;
    sumieff[i]=0;
      sumins[i]=0;


if i=1 then frac=1;
if i=2 then frac=2;
if i=3 then frac=1.67;
if i=4 then frac=1.5;
if i=5 then frac=1.4;
if i=6 then frac=1.33;
if i>=7 then frac=1.33;
dose=dose*frac;

if equation='linear' then dr=min(coeff1+coeff2*dose,1);

if equation='logistic' then dr=exp(coeff1+coeff2*dose)/(1+
exp(coeff1+coeff2*dose));

if equation='loglogistic' then dr=exp(coeff1+coeff2*log(dose))/(1+
exp(coeff1+coeff2*log(dose)));

if i=1 then dreff=0.01;
if i=2 then dreff=0.05;
if i=3 then dreff=0.15;
if i=4 then dreff=0.45;
if i=5 then dreff=0.2;
if i>=6 then dreff=0.05;


r=0;
```

```
p1=dr;
p2=dreff;

  do j=1 to 6;
        a[i,j]=rand("binomial",dr,1);
        q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
          q2=(p2-q1*p1)/(1-p1);
        if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
        else aeff[i,j]=rand("binomial",q2,1);
            if a[i,j]=0 and aeff[i,j]=1 then asucc[i,j]=1;
            else asucc[i,j]=0;
            sumi[i]=sumi[i]+a[i,j];
            sumieff[i]=sumieff[i]+aeff[i,j];
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            onlyrespsdosel[i]=sum(onlyrespsdosel[i],aeff[i,j]/10000);

      respsnodltsdosel[i]=sum(respsnodltsdosel[i],asucc[i,j]/10000);
            sumins[i]=sumins[i]+1;
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=sum(atmtd,1);
            if i>doselevelmtd then abovemtd=sum(abovemtd,1);
            if i<doselevelmtd then belowmtd=sum(belowmtd,1);
            dltc=dltc+a[i,j];
/*
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
*/
            pvaluea[i]=cdf("beta", 0.33, 0.5+sumi[i], 0.5+sumins[i]-
sumi[i]);
          pvalueb[i]=1-cdf("beta", 0.5, 0.5+sumieff[i], 0.5+sumins[i]-
sumieff[i]);
      end;

    if sumi[i]<4 then flagabc[i]=1;
    else flagabc[i]=0;
    jj2=i;

    if flagabc[jj2]=1 then do;

      do j=7 to 14;
        a[i,j]=rand("binomial",dr,1);
        q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
          q2=(p2-q1*p1)/(1-p1);
        if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
        else aeff[i,j]=rand("binomial",q2,1);
            if a[i,j]=0 and aeff[i,j]=1 then asucc[i,j]=1;
            else asucc[i,j]=0;
            sumi[i]=sumi[i]+a[i,j];
            sumieff[i]=sumieff[i]+aeff[i,j];
            sumd=sumd+a[i,j];
```

```
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            onlyrespsdosel[i]=sum(onlyrespsdosel[i],aeff[i,j]/10000);

      respsnodltsdosel[i]=sum(respsnodltsdosel[i],asucc[i,j]/10000);
            sumins[i]=sumins[i]+1;
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=sum(atmtd,1);
            if i>doselevelmtd then abovemtd=sum(abovemtd,1);
            if i<doselevelmtd then belowmtd=sum(belowmtd,1);
            dltc=dltc+a[i,j];
/*
            if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
*/
            pvaluea[i]=cdf("beta", 0.33, 0.5+sumi[i], 0.5+sumins[i]-
sumi[i]);
          pvalueb[i]=1-cdf("beta", 0.5, 0.5+sumieff[i], 0.5+sumins[i]-
sumieff[i]);
      end;


            if ((0<sumi[i]<9 or (sumieff[i]>0 and 0<=sumi[i]<9))) then
do;
        do j=15 to 20;

            a[i,j]=rand("binomial",dr,1);
        q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
          q2=(p2-q1*p1)/(1-p1);
        if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
        else aeff[i,j]=rand("binomial",q2,1);
            if a[i,j]=0 and aeff[i,j]=1 then asucc[i,j]=1;
            else asucc[i,j]=0;
            sumi[i]=sumi[i]+a[i,j];
            sumieff[i]=sumieff[i]+aeff[i,j];
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            onlyrespsdosel[i]=sum(onlyrespsdosel[i],aeff[i,j]/10000);

        respsnodltsdosel[i]=sum(respsnodltsdosel[i],asucc[i,j]/10000);
            sumins[i]=sumins[i]+1;
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=sum(atmtd,1);
            if i>doselevelmtd then abovemtd=sum(abovemtd,1);
            if i<doselevelmtd then belowmtd=sum(belowmtd,1);
            dltc=dltc+a[i,j];
/*
        if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
```

```
*/
            pvaluea[i]=cdf("beta", 0.33, 0.5+sumi[i], 0.5+sumins[i]-
sumi[i]);
          pvalueb[i]=1-cdf("beta", 0.5, 0.5+sumieff[i], 0.5+sumins[i]-
sumieff[i]);
   end;
   end;



  if ((sumi[i]=7 or sumi[i]=8)) then do;

  do j=21 to 26;

            a[i,j]=rand("binomial",dr,1);
        q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
          q2=(p2-q1*p1)/(1-p1);
        if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
        else aeff[i,j]=rand("binomial",q2,1);
            if a[i,j]=0 and aeff[i,j]=1 then asucc[i,j]=1;
            else asucc[i,j]=0;
            sumi[i]=sumi[i]+a[i,j];
            sumieff[i]=sumieff[i]+aeff[i,j];
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            onlyrespsdosel[i]=sum(onlyrespsdosel[i],aeff[i,j]/10000);

      respsnodltsdosel[i]=sum(respsnodltsdosel[i],asucc[i,j]/10000);
            sumins[i]=sumins[i]+1;
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=sum(atmtd,1);
            if i>doselevelmtd then abovemtd=sum(abovemtd,1);
            if i<doselevelmtd then belowmtd=sum(belowmtd,1);
            dltc=dltc+a[i,j];
/*
        if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
*/
            pvaluea[i]=cdf("beta", 0.33, 0.5+sumi[i], 0.5+sumins[i]-
sumi[i]);
          pvalueb[i]=1-cdf("beta", 0.5, 0.5+sumieff[i], 0.5+sumins[i]-
sumieff[i]);
 end;

 if sumi[i]<9 then flagabcdgg[i]=1;
else flagabcdgg[i]=0;
jj27a=i;

if flagabcdgg[jj27a]=1 then do;
 do j=27 to 34;
```

```
            a[i,j]=rand("binomial",dr,1);
        q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
          q2=(p2-q1*p1)/(1-p1);
        if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
        else aeff[i,j]=rand("binomial",q2,1);
            if a[i,j]=0 and aeff[i,j]=1 then asucc[i,j]=1;
            else asucc[i,j]=0;
            sumi[i]=sumi[i]+a[i,j];
            sumieff[i]=sumieff[i]+aeff[i,j];
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            onlyrespsdosel[i]=sum(onlyrespsdosel[i],aeff[i,j]/10000);

      respsnodltsdosel[i]=sum(respsnodltsdosel[i],asucc[i,j]/10000);
            sumins[i]=sumins[i]+1;
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=sum(atmtd,1);
            if i>doselevelmtd then abovemtd=sum(abovemtd,1);
            if i<doselevelmtd then belowmtd=sum(belowmtd,1);
            dltc=dltc+a[i,j];
/*
        if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
*/
            pvaluea[i]=cdf("beta", 0.33, 0.5+sumi[i], 0.5+sumins[i]-
sumi[i]);
          pvalueb[i]=1-cdf("beta", 0.5, 0.5+sumieff[i], 0.5+sumins[i]-
sumieff[i]);
 end;

if sumi[i]<9 then flagabcdhh[i]=1;
else flagabcdhh[i]=0;
jj35a=i;

if flagabcdhh[jj35a]=1 then do;
 do j=35 to 40;

            a[i,j]=rand("binomial",dr,1);
        q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
          q2=(p2-q1*p1)/(1-p1);
        if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
        else aeff[i,j]=rand("binomial",q2,1);
            if a[i,j]=0 and aeff[i,j]=1 then asucc[i,j]=1;
            else asucc[i,j]=0;
            sumi[i]=sumi[i]+a[i,j];
            sumieff[i]=sumieff[i]+aeff[i,j];
            sumd=sumd+a[i,j];
            sump=sump+1;
            peopledosel[i]=sum(peopledosel[i],0.0001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/10000);
            onlyrespsdosel[i]=sum(onlyrespsdosel[i],aeff[i,j]/10000);
```

```
      respsnodltsdosel[i]=sum(respsnodltsdosel[i],asucc[i,j]/10000);
            sumins[i]=sumins[i]+1;
            peoplec=peoplec+1;
            if i=doselevelmtd then atmtd=sum(atmtd,1);
            if i>doselevelmtd then abovemtd=sum(abovemtd,1);
            if i<doselevelmtd then belowmtd=sum(belowmtd,1);
            dltc=dltc+a[i,j];
/*
        if i=doselevelmtd then dltatmtd=dltatmtd+a[i,j];
            if i>doselevelmtd then dltabovemtd=dltabovemtd+a[i,j];
            if i<doselevelmtd then dltbelowmtd=dltbelowmtd+a[i,j];
*/
            pvaluea[i]=cdf("beta", 0.33, 0.5+sumi[i], 0.5+sumins[i]-
sumi[i]);
          pvalueb[i]=1-cdf("beta", 0.5, 0.5+sumieff[i], 0.5+sumins[i]-
sumieff[i]);
end;
end;

end;

end;

end;

end;
end;

/** calculations for ratio of odds of toxicity to odds of efficacy **/
do i2=1 to 10;
if sumieff[i2]*(sumins[i2]-sumi[i2]) not in (0 .) then
ratiocalc=(sumi[i2]*(sumins[i2]-sumieff[i2]))/(sumieff[i2]*(sumins[i2]-
sumi[i2]));
else ratiocalc=.;
ratio[i2]=sum(ratio[i2],ratiocalc);
if sumieff[i2]*(sumins[i2]-sumi[i2]) not in  (0 .) then ration=1;
else ration=.;
ratiof[i2]=sum(ratiof[i2],ration);
end;

/** calculations for satisfying the Bayesian criteria for safety and
efficacy **/
do i3=1 to 10;
if pvaluea[i3]>0.1 and pvalueb[i3]>0.1 then
bayflag[i3]=sum(bayflag[i3],1);
end;

if (pvaluea1<=0.1 or pvalueb1<=0.1) and (pvaluea2<=0.1 or
pvalueb2<=0.1) and (pvaluea3<=0.1 or pvalueb3<=0.1) and (pvaluea4<=0.1
or pvalueb4<=0.1) and (pvaluea5<=0.1 or pvalueb5<=0.1) and
(pvaluea6<=0.1 or pvalueb6<=0.1) and (pvaluea7<=0.1 or pvalueb7<=0.1)
(pvaluea8<=0.1 or pvalueb8<=0.1) (pvaluea9<=0.1 or pvalueb9<=0.1)
(pvaluea10<=0.1 or pvalueb10<=0.1) then bayflag[20]=sum(bayflag[20],1);
```

```
/** calculations for average dlt rate and response rate at each dose
level **/
do i4=1 to 10;
dltratedoselevel[i4]=sum(dltratedoselevel[i4],sumi[i4]/(sumins[i4]*1000
0));
responseratedoselevel[i4]=sum(responseratedoselevel[i4],sumieff[i4]/(su
mins[i4]*10000));
end;

/** calculations for which dose level is chosen in each simulation
based on the value of the utility function **/
do i5=1 to 10;
utilitydltrate[i5]=sumi[i5]/sumins[i5];
utilityresponserate[i5]=sumieff[i5]/sumins[i5];
utility[i5]=utilityresponserate[i5]-0.1*utilitydltrate[i5];
end;

maxutility=max(of utility[*]);

Index[k]=whichN(maxutility,of utility[*]);

dosel[k]=i;
dosemtd[k]=atmtd/peoplec*100;
doseover[k]=abovemtd/peoplec*100;
doseunder[k]=belowmtd/peoplec*100;
/*
dltmtd[k]=dltatmtd/dltc*100;
dltover[k]=dltabovemtd/dltc*100;
dltunder[k]=dltbelowmtd/dltc*100;
*/

dltoverall[k]=dltc/peoplec;
if atmtd ne 0 then mtddltrate[k]=dltatmtd/atmtd;
else mtddltrate[k]=0;
mtdpop[k]=atmtd;
totalpop[k]=peoplec;
totaldlt[k]=dltc;
if (sumins[i]=14 and sumi[i]>=9) then count9=count9+1;

end;

avgd=sumd/10000;
avgp=sump/10000;
ave= mean(of dosel1-dosel10000);
dev=STD(of dosel1-dosel10000);
serror=STDERR(of dosel1-dosel10000);
max= max(of dosel1-dosel10000);
median= median(of dosel1-dosel10000);
pctl_25=pctl(25, of dosel1-dosel10000);
pctl_50=pctl(50, of dosel1-dosel10000);
pctl_75=pctl(75, of dosel1-dosel10000);
avgdosemtd=mean(of dosemtd1-dosemtd10000);
avgdoseover=mean(of doseover1-doseover10000);
avgdoseunder=mean(of doseunder1-doseunder10000);
```

```
/*
avgdltmtd=mean(of dltmtd1-dltmtd10);
avgdltover=mean(of dltover1-dltover10);
avgdltunder=mean(of dltunder1-dltunder10);
*/
avgdltrate=mean(of dltoverall1-dltoverall10000);
avgmtddltrate=mean(of mtddltrate1-mtddltrate10000);
avgmtdpop=median(of mtdpop1-mtdpop10000);
avgtotalpop=median(of totalpop1-totalpop10000);
avgtotalpop1=mean(of totalpop1-totalpop10000);
maxtotalpop=max(of totalpop1-totalpop10000);
mintotalpop=min(of totalpop1-totalpop10000);
avgtotaldlt=median(of totaldlt1-totaldlt10000);
run;

/** print design, highest dose level reached/dose level at which the
trial stops, dose at highest dose level, DLT rate at the highest dose
level, efficacy rate at the highest dose level, factor to multiply
present dose to get the next dose level, max number of dose levels
examined, median, mean, max and min number of patients on trial, median
number of DLTs, average number of patients, DLTs, responders and
responders with no DLTs at each dose level, average DLT rate and
response rate and utility function at each dose level, divide ratio by
ratiof to obtain the average odds ratio of toxicity to efficacy at each
dose level  **/
proc print;
var design i dose dr dreff frac max avgtotalpop avgtotalpop1
maxtotalpop mintotalpop avgtotaldlt peopledosel1-peopledosel10
dltdosel1-dltdosel10 onlyrespsdosel1-onlyrespsdosel10
respsnodltsdosel1-respsnodltsdosel10 dltratedoselevel1-
dltratedoselevel10 responseratedoselevel1-responseratedoselevel10
utility1-utility10 ratio1-ratio10 ratiof1-ratiof10;
run;

/** print for last simulation the number of DLTs at each dose level,
number of responses at each dose level, number of patients at each dose
level, total number of DLTs, total number of patients, p-value for
safety decision criterion and p-value for efficacy decision criterion
at each dose level; also print the number of times each dose level is
chosen based on the Bayesian decision criteria **/
proc print data=simi1;
var sumi1-sumi10 sumieff1-sumieff10 sumins1-sumins10 dltc peoplec
pvaluea1-pvaluea10 pvalueb1-pvalueb10 bayflag1-bayflag10 bayflag20;
run;


%mend alldesigns;

%alldesigns(20+20, logistic);

data simi2;
set simi1;
array v_[*] dosel1-dosel10000;
array counts[10];
```

```
call missing (of counts[*]);
do i = 1 to dim(v_);
     counts[v_[i]] + 1;
end;
if counts1=. then counts1=0;
if counts2=. then counts2=0;
if counts3=. then counts3=0;
if counts4=. then counts4=0;
if counts5=. then counts5=0;
if counts6=. then counts6=0;
if counts7=. then counts7=0;
if counts8=. then counts8=0;
if counts9=. then counts9=0;
if counts10=. then counts10=0;

run;
/** number of times out of 10000 that the simulated trial stops at each
dose level **/
proc print;
var counts1-counts10;
run;

data simi3;
set simi1;
array w_[*] index1-index10000;
array countcounts[10];
call missing (of countcounts[*]);
do i = 1 to dim(w_);
     countcounts[w_[i]] + 1;
end;

run;

/** number of times out of 10000 each dose level is choosen based on
the utility function **/
proc print;
var countcounts1-countcounts10;
run;
```

R code for CRM Design using the R package CRM:

```
# logistic dose-toxicity curve
prior1 <- c(0.15,0.25,0.3,0.45,0.51, 0.56)
true1 <- c(0.01,0.02199,0.06165, 0.2,0.55412,0.8879)
# simulations using model 2 (logistic model)
crmsim(target=0.2,prior=prior1,true=true1,rate=0.1,cycle=21,cohort=20,nsubject=120,nsim=2000,
model=2,a0=1,b=3,jump=FALSE,start.dose=1,seed=777)
```

For the Eff-Tox design, we used the MD Anderson software:

https://biostatistics.mdanderson.org/softwaredownload/SingleSoftware.asp

px?Software_Id=2

For the OBD Isotonic design, we used the R code from Zang et al.:
http://odin.mdacc.tmc.edu/~yyuan/Software/TargetAgent/targetAgentDF.r

For the Phase 2 design calculations for sample size, we used the software EAST with the option "Discrete" for endpoint and "One Sample" for procedure:

/ SAS code for approximate target DLT interval for the 20+20 design **/**

```
proc optmodel;
    var x >= 0 <= 1;

    /* Pr[Bin(40,x)<=8]=0.5 */
    con Mycon1: CDF('BINOM',8,x,40) = 0.5;
    x = 0.2;
    print x (CDF('BINOM',8,x,40));
    solve;
    print x (CDF('BINOM',8,x,40));

    drop Mycon1;

    /* Pr[Bin(20,x)<=6)]=Pr(Bin(20,x)>=9] */
    con Mycon2: CDF('BINOM',6,x,20) = 1 - CDF('BINOM',9,x,20);
    x = 0.35;
    print x (CDF('BINOM',6,x,20)) (1 - CDF('BINOM',9,x,20));
    solve;
    print x (CDF('BINOM',6,x,20)) (1 - CDF('BINOM',9,x,20));
quit;
```

Codes for Chapter 4

SAS code for the extended TEQR design considering efficacy and toxicity:

```
/** The extended TEQR design code below works for a symmetric target
DLT interval (symmetric about the target DLT rate pT). The code will
need modifications for asymmetric target DLT intervals. Also, the
upper limit for toxicity is 0.33 and the lower limit for efficacy is 0.4.
For other thresholds, these numbers need to be changed in the isotonic
regression parts of the code below. **/

data simi2;
call streaminit(1);

array a{10,100};
array aeff{10,100};
array sumi{1000,10};
array sumieff{1000,10};
```

```
array sumitot{1000,10};
array dltr{10};
array effr{10};
array dosel{1000};
array totalpop{1000};
array dr1{10};
array dreff{10};
array peopledosel{10};
array dltdosel{10};

/** defining the unacceptable dlt probability, the target dlt rate,
the equivalence range for the target DLT rate, the max sample size
at the MTD, the cohort size and the correlation coefficient
(correlation between the true toxicity and true efficacy rates) */
maxdlt=0.34;
target=0.2;
trange=0.05;
maxmtdss=50;
cohortsize=5;
r=0;

do k=1 to 1000;

/* true underlying DLT and response rates */
dr1[1]=0.01;
dr1[2] =0.02199;
dr1[3]=0.06165;
dr1[4]=0.2;
dr1[5]=0.55412;
dr1[6]=0.88790;
dr1[7]=0.98936;
dr1[8]=0.99959;
dr1[9]=0.99999;
dr1[10]=1;


dreff[1]=0.1;
dreff[2]=0.3;
dreff[3]=0.4;
dreff[4]=0.45;
dreff[5]=0.55;
dreff[6]=0.6;
dreff[7]=0.65;


do i1=1 to 10;
do j1=1 to maxmtdss;
        a[i1,j1]=.;
        aeff[i1,j1]=.;
           sumi[k,1]=0;
           sumieff[k,1]=0;
```

```
            sumitot[k,1]=0;
            dltr[i1]=0;
            effr[i1]=0;
end;
end;


i=2; /** start at dose level 2 to allow de-escalation **/
sumi[k,2]=0;
sumieff[k,2]=0;
sumitot[k,2]=0;
peoplec=0;

do j=1 to cohortsize;
            a[i,j]=rand('Bernoulli', dr1[i]);
            sumi[k,i]=sumi[k,i]+a[i,j];
            sumitot[k,i]=sumitot[k,i]+1;

            p1=dr1[i];
         p2=dreff[i];
            q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
           q2=(p2-q1*p1)/(1-p1);
         if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
         else aeff[i,j]=rand("binomial",q2,1);
            sumieff[k,i]=sumieff[k,i]+aeff[i,j];

            dltr[i]=sumi[k,i]/sumitot[k,i];
            effr[i]=sumieff[k,i]/sumitot[k,i];

            peoplec=peoplec+1;
            peopledosel[i]=sum(peopledosel[i],0.001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/1000);
end;

/** dosing decisions of the TEQR design (escalate, stay at the same
dose, de-escalate) performed until the maximum MTD sample size is
reached or we need to de-escalate beyond dose level 1 **/
 do until ((sumsum=maxmtdss and dltrate<maxdlt) or i=0);


    if 0<=dltr[i]<(target-trange) and dltr[i+1]<maxdlt then do;
    i=i+1;
      range=sum(sumitot[k,i],0);
      do j=range+1 to range+cohortsize;
            a[i,j]=rand('Bernoulli', dr1[i]);
            sumi[k,i]=sum(sumi[k,i],a[i,j]);
            sumitot[k,i]=sum(sumitot[k,i],1);
            sumsum=sumitot[k,i];

            p1=dr1[i];
        p2=dreff[i];
```

```
            q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
          q2=(p2-q1*p1)/(1-p1);
        if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
        else aeff[i,j]=rand("binomial",q2,1);
            sumieff[k,i]=sum(sumieff[k,i],aeff[i,j]);

            dltr[i]=sumi[k,i]/sumitot[k,i];
            dltrate=dltr[i];
            effr[i]=sumieff[k,i]/sumitot[k,i];
            effrate=effr[i];

            peoplec=peoplec+1;
            peopledosel[i]=sum(peopledosel[i],0.001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/1000);
      end;
    end;
 else if 0<=dltr[i]<(target-trange) and dltr[i+1]>=maxdlt then do;
      range=sum(sumitot[k,i],0);
      do j=range+1 to range+cohortsize;
            a[i,j]=rand('Bernoulli', dr1[i]);
            sumi[k,i]=sum(sumi[k,i],a[i,j]);
            sumitot[k,i]=sum(sumitot[k,i],1);
            sumsum=sumitot[k,i];

            p1=dr1[i];
        p2=dreff[i];
            q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
          q2=(p2-q1*p1)/(1-p1);
        if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
        else aeff[i,j]=rand("binomial",q2,1);
            sumieff[k,i]=sum(sumieff[k,i],aeff[i,j]);

            dltr[i]=sumi[k,i]/sumitot[k,i];
            dltrate=dltr[i];
            effr[i]=sumieff[k,i]/sumitot[k,i];
            effrate=effr[i];

            peoplec=peoplec+1;
            peopledosel[i]=sum(peopledosel[i],0.001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/1000);
      end;
    end;

else if (target-trange)<=dltr[i]<=(target+trange) then do;
range=sum(sumitot[k,i],0);
     do j=range+1 to range+cohortsize;
            a[i,j]=rand('Bernoulli', dr1[i]);
            sumi[k,i]=sum(sumi[k,i],a[i,j]);
            sumitot[k,i]=sum(sumitot[k,i],1);
            sumsum=sumitot[k,i];
```

```
            p1=dr1[i];
          p2=dreff[i];
            q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
           q2=(p2-q1*p1)/(1-p1);
          if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
          else aeff[i,j]=rand("binomial",q2,1);
            sumieff[k,i]=sum(sumieff[k,i],aeff[i,j]);

            dltr[i]=sumi[k,i]/sumitot[k,i];
            dltrate=dltr[i];
            effr[i]=sumieff[k,i]/sumitot[k,i];
            effrate=effr[i];

            peoplec=peoplec+1;
            peopledosel[i]=sum(peopledosel[i],0.001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/1000);
        end;
      end;

        else if (target+trange)<dltr[i]<=1 then do;
      i=i-1;
        if i > 0 then do;
        range=sum(sumitot[k,i],0);
        do j=range+1 to range+cohortsize;
            a[i,j]=rand('Bernoulli', dr1[i]);
            sumi[k,i]=sum(sumi[k,i],a[i,j]);
            sumitot[k,i]=sum(sumitot[k,i],1);
            sumsum=sumitot[k,i];

            p1=dr1[i];
          p2=dreff[i];
            q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
           q2=(p2-q1*p1)/(1-p1);
          if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
          else aeff[i,j]=rand("binomial",q2,1);
            sumieff[k,i]=sum(sumieff[k,i],aeff[i,j]);

            dltr[i]=sumi[k,i]/sumitot[k,i];
            dltrate=dltr[i];
            effr[i]=sumieff[k,i]/sumitot[k,i];
            effrate=effr[i];

            peoplec=peoplec+1;
            peopledosel[i]=sum(peopledosel[i],0.001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/1000);
        end;
      end;
        end;

end;
```

```
if i ne 0 then dosel[k]=i;
if i eq 0 then dosel[k]=20;

totalpop[k]=peoplec;

end;

maxdose= max(of dosel1-dosel1000);
medianpeop= median(of totalpop1-totalpop1000);
run;

/* print median number of patients on trial, avg DLT rate and
response rate in each dose level, average number of patients and
DLTs in each dose level */
proc print;
var medianpeop dltr1-dltr10 effr1-effr10 peopledosel1-peopledosel10
dltdosel1-dltdosel10;
run;


data simi3;
set simi2 (keep=dosel1-dosel1000);

array v_[*] dosel1-dosel1000;
array counts[20];
call missing (of counts[*]);
do i = 1 to dim(v_);
     counts[v_[i]] + 1;
end;

run;

/** number of times out of 1000 that each dose level is selected for
safety before isotonic regression **/
proc print;
var counts1-counts20;
run;

/* isotonic regression for observed DLT rates */

%macro doit(k1);

%do jj=1  %to &k1;

%global max_els2;

%let onedim=Yes; *Yes/No;


%let num1=%eval((&jj-1)*10+1);
%let num2=%eval((&jj-1)*10+2);
```

```
%let num3=%eval((&jj-1)*10+3);
%let num4=%eval((&jj-1)*10+4);
%let num5=%eval((&jj-1)*10+5);
%let num6=%eval((&jj-1)*10+6);
%let num7=%eval((&jj-1)*10+7);
%let num8=%eval((&jj-1)*10+8);
%let num9=%eval((&jj-1)*10+9);
%let num10=%eval((&jj-1)*10+10);


proc iml;
use simi2 ;

read all var{sumi&num1 sumi&num2 sumi&num3 sumi&num4 sumi&num5
sumi&num6 sumi&num7 sumi&num8 sumi&num9 sumi&num10} into x;

read all var{sumitot&num1 sumitot&num2 sumitot&num3 sumitot&num4
sumitot&num5 sumitot&num6 sumitot&num7 sumitot&num8 sumitot&num9
sumitot&num10} into w;

mi=missing(x);
rowmi = mi[,+];
rowmi1=10-rowmi;

call symput('rowmi2',char(rowmi1));

/*
print mi rowmi rowmi1;
*/

create one1 from w;
append from w;
create one2 from x;
append from x;
quit;
run;

%let max_els2=&rowmi2; *Number rows/columns;

%macro assign(v,out);

data &out;
set &out;
array cols[*] col1-col&max_els2;
array s[*] &v.1-&v.&max_els2;
do i=1 to &max_els2;
%if &v=w %then %do; s[i]=max(0.0001,cols[i]); %end;
%else %do; s[i]=cols[i]; %end;
end;
keep &v:;
run;
```

```
%mend;
%assign(v=w,out=one1);
%assign(v=x,out=one2);

data one;
merge one1 one2;
array ys[*] y1-y&max_els2;
array ws[*] w1-w&max_els2;
array xs[*] x1-x&max_els2;
array zs[*] z1-z&max_els2;
array ss[*] s1-s&max_els2;
do i=1 to &max_els2;
if ws[i]>0 then ys[i]=xs[i]/ws[i];
else ys[i]=0;
zs[i]=ys[i];
ss[i]=0;
end;
drop i x:;
run;

proc datasets;
delete one1 one2;
run;

/*
proc print n width=min;
title 'orig';
var w1-w&max_els2 y1-y&max_els2;
run;
*/

%let flag=0;
%let n=0;
%do %until(&flag=1 or (&onedim=Yes and &n=1));
%let n=%eval(&n+1);

%macro rename(v,out);

data &out;
retain &v.1-&v.&max_els2;
keep &v.1-&v.&max_els2;
set one;
run;

%mend;
%rename(v=y,out=y);
%rename(v=w,out=w);
%rename(v=z,out=z);
%rename(v=s,out=s);
```

```
%if &onedim^=Yes %then %do;
%if &n^=1 %then %do;
proc iml;
use y; read all into one1;
use w; read all into one2;
use z; read all into one3;
use s; read all into one4;
new=T(one1);
new2=T(one2);
new3=T(one3);
new4=T(one4);
create one1 from new;
append from new;
create one2 from new2;
append from new2;
create one3 from new3;
append from new3;
create one4 from new4;
append from new4;
quit;
run;

%assign(v=y,out=one1);
%assign(v=w,out=one2);
%assign(v=z,out=one3);
%assign(v=s,out=one4);

data one;
merge one1 one2 one3 one4;
run;

proc datasets;
delete one1 one2 one3 one4 y w z s;
run;
%end;
%end;

data one;
set one;

%macro pav(max_els=,array=,weights=);

array _yo_[&max_els] yo1-yo&max_els;
array _w_[&max_els] w1-w&max_els;
array _y_[&max_els] y1-y&max_els;
array _z_[&max_els] z1-z&max_els;
array _s_[&max_els] s1-s&max_els;

do i=1 to &max_els;
_yo_[i]=_y_[i];
end;
```

```
%if &n=2 %then %do;
do i=1 to &max_els;
_y_[i]=_z_[i];
end;
%end;
%else %if &n>=3 %then %do;
do i=1 to &max_els;
*_y_[i]=max(0,_z_[i]+_s_[i]);
_y_[i]=_z_[i]+_s_[i];
end;
%end;

%global index ;
%if %quote(&index)=
%then
%let index=1 ;
%else
   %let index=%eval(&index+1) ;
%let pooled = _pool&index._ ;
%let parray = _parr&index._ ;
%let pwghts = _pwgt&index._ ;

array &pooled {&max_els} _TEMPORARY_ ;
array &parray {&max_els} _TEMPORARY_ ;
array &pwghts {&max_els} _TEMPORARY_ ;

if dim(&array) = 1 then Go to epav&index ;

do _pav_j_ = 1 to dim(&array) ;
&pooled(_pav_j_) = 0 ;
&parray(_pav_j_) = 0 ;
&pwghts(_pav_j_) = 0 ;
end ;

&parray(1) = &array(1) ;
&pwghts(1) = &weights(1) ;
_pav_j_ = 1 ;

do _pav_i_ = 2 to dim(&array) ;

  /* if ajacent violated, then pool */
  if (&parray(_pav_j_) > &array(_pav_i_)) then do ;
     _plwght_ = &pwghts(_pav_j_) + &weights(_pav_i_) ;
    _plval_ = ((&parray(_pav_j_)*&pwghts(_pav_j_)) +
              (&array(_pav_i_)*&weights(_pav_i_))) / _plwght_ ;

    &pooled(_pav_i_) = 1 ;
    flag=0; *added;
    if _pav_j_ > 1 then do ;
       _pav_j_ = _pav_j_ - 1 ;
```

```
       _pav_jj_ = _pav_i_ ;
       do while (flag=0 & ( &parray(_pav_j_)>_plval_) & (_pav_i_ >=
1)); *added flag=0;
       put &parray(_pav_j_)= _pav_j_= _plval_=;
         _tplval_ = _plval_ ;
         _tplwgt_ = _plwght_ ; /* tplwgt misspelled */
          do until (not &pooled(_pav_jj_)) ;
              _pav_jj_ = _pav_jj_ - 1 ;
           put _pav_jj_= &pooled(_pav_jj_)=;
              end ; /* do until */

          _plwght_ = &pwghts(_pav_j_) + _tplwgt_ ;
          _plval_  = ((&parray(_pav_j_) * &pwghts(_pav_j_)) +
                      (_tplval_*_tplwgt_)) / _plwght_ ;
         &pooled(_pav_jj_) = 1 ;
         _pav_j_ = _pav_j_ -1 ;
         if _pav_j_=0 then do; flag=1; _pav_j_=1; end; *added;
         end ; /* do while */
      _pav_j_ = _pav_j_ + 1 ;
      end ; /* if _pav_j_ > 1 */
      if flag=1 then do; *added;
         _pav_j_=_pav_j_-1; *added;
         do i=1 to _pav_i_;  *added*;
         &parray(i) = _plval_ ; *added;
         &pwghts(i) = _plwght_ ; *added;
         end;  *added*;
         end;  *added*;
     else do; *added;
     &parray(_pav_j_) = _plval_ ;
     &pwghts(_pav_j_) = _plwght_ ;
     end; *added;
     end ; /* if (&parray... */

     else do ;
         _pav_j_ = _pav_j_ + 1 ;
          &parray(_pav_j_) = &array(_pav_i_) ;
          &pwghts(_pav_j_) = &weights(_pav_i_) ;
          end ;
     end ; /* _pav_i_=2 to dim(array) */

  &array(1) = &parray(1) ;
 _pav_j_=1 ;
 _pav_jj_=1 ; /*left off underscore*/

 /* put the pooled data back into the original array */
 do _pav_j_ = 2 to dim(&array) ;
   if ^&pooled(_pav_j_) then _pav_jj_ = _pav_jj_ + 1 ;
   &array(_pav_j_) = &parray(_pav_jj_) ;
 end ;

 Epav&index:
```

```
 drop _pav_j_ _pav_i_ _pav_jj_ _plval_ _plwght_ _tplval_ _tplwgt_ ;

          ok=0;
          do i=1 to &max_els;
          %if &n=1 %then %do;
          _s_[i]=_y_[i]-_z_[i]; %end;
          %else %if &n>=2 %then %do;
          _s_[i]=_y_[i]-_z_[i]-_s_[i]; %end;
          if abs(_yo_[i]-_y_[i])>.0001 then ok=ok+1;
          end;
              run;

              %mend;

%pav(max_els=&max_els2,array=_y_,weights=_w_);

data one;
set one;
index2=1;
run;

data one;
set one end=lastrec;
by index2;
retain ok2;
if first.index2 then ok2=ok;
else ok2=ok2+ok;
if lastrec and ((mod(&n,2)^=0 and ok2=0) or &n=199) then call
symput('flag','1');
*if lastrec and (&n=51) then call symput('flag','1');

/** highest dose level with DLT rate <=0.33 after isotonic
regression **/
if y1>0.33 then mtdiso=0;
if .<y1<=0.33 then mtdiso=1;
if .<y2<=0.33 then mtdiso=2;
if .<y3<=0.33 then mtdiso=3;
if .<y4<=0.33 then mtdiso=4;
if .<y5<=0.33 then mtdiso=5;
if .<y6<=0.33 then mtdiso=6;
if .<y7<=0.33 then mtdiso=7;
if .<y8<=0.33 then mtdiso=8;
if .<y9<=0.33 then mtdiso=9;
if .<y10<=0.33 then mtdiso=10;

drop i index2 ok:;
run;

/*
proc print; var  w1-w&max_els2 y1-y&max_els2; title "Iter &n After
Iso"; run;
```

```
*/

data finalmtdiso&jj;
set one;
keep mtdiso;
run;

%end;

%end;
%mend;


%doit(1000);

data finaliso;
set finalmtdiso1 - finalmtdiso1000;
run;

/** number of times out of 1000 that each dose level is selected for
safety after isotonic regression (highest dose level at which the
smoothed DLT rate is <=0.33) **/
proc freq;
tables mtdiso;
run;

***********;

/* isotonic regression for observed response rates */

%macro doiteff(k1);

%do jj=1  %to &k1;

%global max_els2;

%let onedim=Yes; *Yes/No;


%let num1=%eval((&jj-1)*10+1);
%let num2=%eval((&jj-1)*10+2);
%let num3=%eval((&jj-1)*10+3);
%let num4=%eval((&jj-1)*10+4);
%let num5=%eval((&jj-1)*10+5);
%let num6=%eval((&jj-1)*10+6);
%let num7=%eval((&jj-1)*10+7);
%let num8=%eval((&jj-1)*10+8);
%let num9=%eval((&jj-1)*10+9);
%let num10=%eval((&jj-1)*10+10);
```

```
proc iml;
use simi2 ;

read all var{sumieff&num1 sumieff&num2 sumieff&num3 sumieff&num4
sumieff&num5 sumieff&num6 sumieff&num7 sumieff&num8 sumieff&num9
sumieff&num10} into x;

read all var{sumitot&num1 sumitot&num2 sumitot&num3 sumitot&num4
sumitot&num5 sumitot&num6 sumitot&num7 sumitot&num8 sumitot&num9
sumitot&num10} into w;

mi=missing(x);
rowmi = mi[,+];
rowmi1=10-rowmi;

call symput('rowmi2',char(rowmi1));

/*
print mi rowmi rowmi1;
*/

create one1 from w;
append from w;
create one2 from x;
append from x;
quit;
run;

%let max_els2=&rowmi2; *Number rows/columns;

%macro assign(v,out);

data &out;
set &out;
array cols[*] col1-col&max_els2;
array s[*] &v.1-&v.&max_els2;
do i=1 to &max_els2;
%if &v=w %then %do; s[i]=max(0.0001,cols[i]); %end;
%else %do; s[i]=cols[i]; %end;
end;
keep &v:;
run;

%mend;
%assign(v=w,out=one1);
%assign(v=x,out=one2);

data one;
merge one1 one2;
array ys[*] y1-y&max_els2;
array ws[*] w1-w&max_els2;
```

```
array xs[*] x1-x&max_els2;
array zs[*] z1-z&max_els2;
array ss[*] s1-s&max_els2;
do i=1 to &max_els2;
if ws[i]>0 then ys[i]=xs[i]/ws[i];
else ys[i]=0;
zs[i]=ys[i];
ss[i]=0;
end;
drop i x:;
run;

proc datasets;
delete one1 one2;
run;

/*
proc print n width=min;
title 'orig';
var w1-w&max_els2 y1-y&max_els2;
run;
*/

%let flag=0;
%let n=0;
%do %until(&flag=1 or (&onedim=Yes and &n=1));
%let n=%eval(&n+1);

%macro rename(v,out);

data &out;
retain &v.1-&v.&max_els2;
keep &v.1-&v.&max_els2;
set one;
run;

%mend;
%rename(v=y,out=y);
%rename(v=w,out=w);
%rename(v=z,out=z);
%rename(v=s,out=s);

%if &onedim^=Yes %then %do;
%if &n^=1 %then %do;
proc iml;
use y; read all into one1;
use w; read all into one2;
use z; read all into one3;
use s; read all into one4;
new=T(one1);
new2=T(one2);
```

```
new3=T(one3);
new4=T(one4);
create one1 from new;
append from new;
create one2 from new2;
append from new2;
create one3 from new3;
append from new3;
create one4 from new4;
append from new4;
quit;
run;

%assign(v=y,out=one1);
%assign(v=w,out=one2);
%assign(v=z,out=one3);
%assign(v=s,out=one4);

data one;
merge one1 one2 one3 one4;
run;

proc datasets;
delete one1 one2 one3 one4 y w z s;
run;
%end;
%end;

data one;
set one;

%macro pav(max_els=,array=,weights=);

array _yo_[&max_els] yo1-yo&max_els;
array _w_[&max_els] w1-w&max_els;
array _y_[&max_els] y1-y&max_els;
array _z_[&max_els] z1-z&max_els;
array _s_[&max_els] s1-s&max_els;

do i=1 to &max_els;
_yo_[i]=_y_[i];
end;

%if &n=2 %then %do;
do i=1 to &max_els;
_y_[i]=_z_[i];
end;
%end;
%else %if &n>=3 %then %do;
do i=1 to &max_els;
*_y_[i]=max(0,_z_[i]+_s_[i]);
```

```
_y_[i]=_z_[i]+_s_[i];
end;
%end;

%global index ;
%if %quote(&index)=
%then
%let index=1 ;
%else
   %let index=%eval(&index+1) ;
%let pooled = _pool&index._ ;
%let parray = _parr&index._ ;
%let pwghts = _pwgt&index._ ;

array &pooled {&max_els} _TEMPORARY_ ;
array &parray {&max_els} _TEMPORARY_ ;
array &pwghts {&max_els} _TEMPORARY_ ;

if dim(&array) = 1 then Go to epav&index ;

do _pav_j_ = 1 to dim(&array) ;
&pooled(_pav_j_) = 0 ;
&parray(_pav_j_) = 0 ;
&pwghts(_pav_j_) = 0 ;
end ;

&parray(1) = &array(1) ;
&pwghts(1) = &weights(1) ;
_pav_j_ = 1 ;

do _pav_i_ = 2 to dim(&array) ;

  /* if ajacent violated, then pool */
  if (&parray(_pav_j_) > &array(_pav_i_)) then do ;
     _plwght_ = &pwghts(_pav_j_) + &weights(_pav_i_) ;
    _plval_ = ((&parray(_pav_j_)*&pwghts(_pav_j_)) +
              (&array(_pav_i_)*&weights(_pav_i_))) / _plwght_ ;

    &pooled(_pav_i_) = 1 ;
    flag=0; *added;
    if _pav_j_ > 1 then do ;
       _pav_j_ = _pav_j_ - 1 ;
       _pav_jj_ = _pav_i_ ;
       do while (flag=0 & ( &parray(_pav_j_)>_plval_) & (_pav_i_ >=
1)); *added flag=0;
       put &parray(_pav_j_)= _pav_j_= _plval_=;
         _tplval_ = _plval_ ;
         _tplwgt_ = _plwght_ ; /* tplwgt misspelled */
          do until (not &pooled(_pav_jj_)) ;
              _pav_jj_ = _pav_jj_ - 1 ;
           put _pav_jj_= &pooled(_pav_jj_)=;
```

```
            end ; /* do until */

          _plwght_ = &pwghts(_pav_j_) + _tplwgt_ ;
          _plval_  = ((&parray(_pav_j_) * &pwghts(_pav_j_)) +
                      (_tplval_*_tplwgt_)) / _plwght_ ;
         &pooled(_pav_jj_) = 1 ;
         _pav_j_ = _pav_j_ -1 ;
         if _pav_j_=0 then do; flag=1; _pav_j_=1; end; *added;
         end ; /* do while */
       _pav_j_ = _pav_j_ + 1 ;
       end ; /* if _pav_j_ > 1 */
       if flag=1 then do; *added;
          _pav_j_=_pav_j_-1; *added;
          do i=1 to _pav_i_;  *added*;
          &parray(i) = _plval_ ; *added;
          &pwghts(i) = _plwght_ ; *added;
          end;  *added*;
          end;  *added*;
      else do; *added;
      &parray(_pav_j_) = _plval_ ;
      &pwghts(_pav_j_) = _plwght_ ;
      end; *added;
      end ; /* if (&parray... */

      else do ;
          _pav_j_ = _pav_j_ + 1 ;
           &parray(_pav_j_) = &array(_pav_i_) ;
           &pwghts(_pav_j_) = &weights(_pav_i_) ;
           end ;
      end ; /* _pav_i_=2 to dim(array) */

 &array(1) = &parray(1) ;
_pav_j_=1 ;
_pav_jj_=1 ; /*left off underscore*/

/* put the pooled data back into the original array */
do _pav_j_ = 2 to dim(&array) ;
  if ^&pooled(_pav_j_) then _pav_jj_ = _pav_jj_ + 1 ;
  &array(_pav_j_) = &parray(_pav_jj_) ;
end ;

Epav&index:
drop _pav_j_ _pav_i_ _pav_jj_ _plval_ _plwght_ _tplval_ _tplwgt_ ;

          ok=0;
          do i=1 to &max_els;
          %if &n=1 %then %do;
          _s_[i]=_y_[i]-_z_[i]; %end;
          %else %if &n>=2 %then %do;
          _s_[i]=_y_[i]-_z_[i]-_s_[i]; %end;
          if abs(_yo_[i]-_y_[i])>.0001 then ok=ok+1;
```

```
          end;
              run;

              %mend;

%pav(max_els=&max_els2,array=_y_,weights=_w_);

data one;
set one;
index2=1;
run;

data one;
set one end=lastrec;
by index2;
retain ok2;
if first.index2 then ok2=ok;
else ok2=ok2+ok;
if lastrec and ((mod(&n,2)^=0 and ok2=0) or &n=199) then call
symput('flag','1');
*if lastrec and (&n=51) then call symput('flag','1');

/** lowest dose level with response rate >=0.4 after isotonic
regression **/
effiso=0;
if y1>=0.4 then effiso=1;
else if y2>=0.4 then effiso=2;
else if y3>=0.4 then effiso=3;
else if y4>=0.4 then effiso=4;
else if y5>=0.4 then effiso=5;
else if y6>=0.4 then effiso=6;
else if y7>=0.4 then effiso=7;
else if y8>=0.4 then effiso=8;
else if y9>=0.4 then effiso=9;
else if y10>=0.4 then effiso=10;

drop i index2 ok:;
run;

/*
proc print; var  w1-w&max_els2 y1-y&max_els2; title "Iter &n After
Iso"; run;
*/

data finaleffiso&jj;
set one;
keep effiso;
run;

%end;
```

```
%end;
%mend;

%doiteff(1000);

data finaliso1;
set finaleffiso1 - finaleffiso1000;
run;

/** number of times out of 1000 that each dose level is selected for
efficacy after isotonic regression (lowest dose level at which the
smoothed response rate is >=0.4) **/
proc freq;
tables effiso;
run;

*****************;
/** code for number of times out of 1000 that each dose is selected as
optimal dose when the dlt rates are monotonically increasing and the
response rates are monotonically increasing or monotonically non-
decreasing **/

data finaliso111;
set finaliso;
retain a;
a=sum(a,1);
run;

data finaliso112;
set finaliso1;
retain a;
a=sum(a,1);
run;

data teqrisoiso;
merge finaliso111 finaliso112;
by a;
if effiso ne 0 and effiso<=mtdiso then optimal=mtdiso;
else optimal=0;
run;

proc freq;
tables optimal;
run;
```

SAS code for the extended mTPI design considering efficacy and toxicity:

```
/** The extended mTPI design code below works for a symmetric target
DLT interval (symmetric about the target DLT rate pT). The code will
need modifications for asymmetric target DLT intervals. Also, the
```

```
upper limit for toxicity is 0.33 and the lower limit for efficacy is 0.4.
For other thresholds, these numbers need to be changed in the isotonic
regression parts of the code below. **/

data simi2;
call streaminit(1);

array a{10,100};
array aeff{10,100};
array sumi{1000,10};
array sumieff{1000,10};
array sumitot{1000,10};
array dltr{10};
array effr{10};
array calc1{10};
array calc2{10};
array calc3{10};
array calc4{10};
array dosel{1000};
array totalpop{1000};
array dr1{10};
array dreff{10};
array peopledosel{10};
array dltdosel{10};

/** defining the target dlt rate, the equivalence range for the
target DLT rate, the max sample size, the cohort size and the
correlation coefficient (correlation between the true toxicity and
true efficacy rates) */

target=0.2;
trange=0.05;
maxss=50;
cohortsize=5;
r=0;

do k=1 to 1000;

/** true underlying DLT and response rates **/
dr1[1]=0.01;
dr1[2] =0.02199;
dr1[3]=0.06165;
dr1[4]=0.2;
dr1[5]=0.55412;
dr1[6]=0.88790;
dr1[7]=0.98936;
dr1[8]=0.99959;
dr1[9]=0.99999;
dr1[10]=1;

dreff[1]=0.1;
```

```
dreff[2]=0.3;
dreff[3]=0.4;
dreff[4]=0.45;
dreff[5]=0.55;
dreff[6]=0.6;
dreff[7]=0.65;


do i1=1 to 10;
do j1=1 to maxss;
        a[i1,j1]=.;
        aeff[i1,j1]=.;
            sumi[k,1]=0;
            sumieff[k,1]=0;
            sumitot[k,1]=0;
            dltr[i1]=0;
            effr[i1]=0;
            calc1[i1]=0;
            calc2[i1]=0;
            calc3[i1]=0;
            calc4[i1]=0;
end;
end;


i=2; /** start at dose level 2 to allow de-escalation **/
sumi[k,2]=0;
sumieff[k,2]=0;
sumitot[k,2]=0;
peoplec=0;

do j=1 to cohortsize;
            if ranuni(1)<dr1[i] then a[i,j]=1; else a[i,j]=0;
/*a[i,j]=rand('Bernoulli', dr1[i]);*/
            sumi[k,i]=sumi[k,i]+a[i,j];
            sumitot[k,i]=sumitot[k,i]+1;

            p1=dr1[i];
        p2=dreff[i];
            q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
          q2=(p2-q1*p1)/(1-p1);
        if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
        else aeff[i,j]=rand("binomial",q2,1);
            sumieff[k,i]=sumieff[k,i]+aeff[i,j];

            dltr[i]=sumi[k,i]/sumitot[k,i];
            effr[i]=sumieff[k,i]/sumitot[k,i];

            calc1[i]= (1-cdf("beta", target+trange, sumi[k,i]+1,
sumitot[k,i]-sumi[k,i]+1))/(1-target-trange);
```

```
            calc2[i]=  (cdf("beta", target+trange, sumi[k,i]+1,
sumitot[k,i]-sumi[k,i]+1) - cdf("beta", target-trange, sumi[k,i]+1,
sumitot[k,i]-sumi[k,i]+1))/(trange+trange);
            calc3[i]= (cdf("beta", target-trange, sumi[k,i]+1,
sumitot[k,i]-sumi[k,i]+1)/(target-trange));
            calc4[i]= 1-cdf("beta", target, sumi[k,i]+1,
sumitot[k,i]-sumi[k,i]+1);

            peoplec=peoplec+1;
            peopledosel[i]=sum(peopledosel[i],0.001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/1000);
end;


/** dosing decisions of the mTPI design (escalate, stay at the same
dose, de-escalate) performed until the maximum sample size is
reached or we need to de-escalate beyond dose level 1 **/
 do until ((peoplec=maxss) or i=0);


    if calc3[i]>=calc2[i] and calc3[i]>calc1[i] and calc4[i+1]<=0.95
then do;
    i=i+1;
      range=sum(sumitot[k,i],0);
      do j=range+1 to range+cohortsize;
            if ranuni(1)<dr1[i] then a[i,j]=1; else a[i,j]=0;
/*a[i,j]=rand('Bernoulli', dr1[i]);*/
            sumi[k,i]=sum(sumi[k,i],a[i,j]);
            sumitot[k,i]=sum(sumitot[k,i],1);
            sumsum=sumitot[k,i];

            p1=dr1[i];
         p2=dreff[i];
            q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
           q2=(p2-q1*p1)/(1-p1);
         if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
         else aeff[i,j]=rand("binomial",q2,1);
            sumieff[k,i]=sum(sumieff[k,i],aeff[i,j]);

            dltr[i]=sumi[k,i]/sumitot[k,i];
            dltrate=dltr[i];
            effr[i]=sumieff[k,i]/sumitot[k,i];
            effrate=effr[i];

            calc1[i]= (1-cdf("beta", target+trange, sum(sumi[k,i],1),
sum(sumitot[k,i],-sumi[k,i])+1))/(1-target-trange);
            calc2[i]=  (cdf("beta", target+trange, sum(sumi[k,i],1),
sum(sumitot[k,i],-sumi[k,i])+1) - cdf("beta", target-trange,
sum(sumi[k,i],1), sum(sumitot[k,i],-sumi[k,i])+1))/(trange+trange);
            calc3[i]= (cdf("beta", target-trange, sum(sumi[k,i],1),
sum(sumitot[k,i],-sumi[k,i])+1)/(target-trange));
```

```
            calc4[i]= 1-cdf("beta", target, sum(sumi[k,i],1),
sum(sumitot[k,i],-sumi[k,i])+1);

            peoplec=peoplec+1;
            peopledosel[i]=sum(peopledosel[i],0.001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/1000);
      end;
    end;

  else if calc3[i]>=calc2[i] and calc3[i]>calc1[i] and
calc4[i+1]>0.95 then do;
      range=sum(sumitot[k,i],0);
      do j=range+1 to range+cohortsize;
            if ranuni(1)<dr1[i] then a[i,j]=1; else a[i,j]=0;
/*a[i,j]=rand('Bernoulli', dr1[i]);*/
            sumi[k,i]=sum(sumi[k,i],a[i,j]);
            sumitot[k,i]=sum(sumitot[k,i],1);
            sumsum=sumitot[k,i];

            p1=dr1[i];
        p2=dreff[i];
            q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
          q2=(p2-q1*p1)/(1-p1);
        if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
        else aeff[i,j]=rand("binomial",q2,1);
            sumieff[k,i]=sum(sumieff[k,i],aeff[i,j]);

            dltr[i]=sumi[k,i]/sumitot[k,i];
            dltrate=dltr[i];
            effr[i]=sumieff[k,i]/sumitot[k,i];
            effrate=effr[i];

            calc1[i]= (1-cdf("beta", target+trange, sum(sumi[k,i],1),
sum(sumitot[k,i],-sumi[k,i])+1))/(1-target-trange);
            calc2[i]=  (cdf("beta", target+trange, sum(sumi[k,i],1),
sum(sumitot[k,i],-sumi[k,i])+1) - cdf("beta", target-trange,
sum(sumi[k,i],1), sum(sumitot[k,i],-sumi[k,i])+1))/(trange+trange);
            calc3[i]= (cdf("beta", target-trange, sum(sumi[k,i],1),
sum(sumitot[k,i],-sumi[k,i])+1)/(target-trange));
            calc4[i]= 1-cdf("beta", target, sum(sumi[k,i],1),
sum(sumitot[k,i],-sumi[k,i])+1);

            peoplec=peoplec+1;
            peopledosel[i]=sum(peopledosel[i],0.001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/1000);
      end;
    end;

 else if calc2[i]>=calc1[i] and calc2[i]>calc3[i] then do;
 range=sum(sumitot[k,i],0);
      do j=range+1 to range+cohortsize;
```

```
            if ranuni(1)<dr1[i] then a[i,j]=1; else a[i,j]=0;
/*a[i,j]=rand('Bernoulli', dr1[i]);*/
            sumi[k,i]=sum(sumi[k,i],a[i,j]);
            sumitot[k,i]=sum(sumitot[k,i],1);
            sumsum=sumitot[k,i];

            p1=dr1[i];
         p2=dreff[i];
            q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
           q2=(p2-q1*p1)/(1-p1);
         if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
         else aeff[i,j]=rand("binomial",q2,1);
            sumieff[k,i]=sum(sumieff[k,i],aeff[i,j]);

            dltr[i]=sumi[k,i]/sumitot[k,i];
            dltrate=dltr[i];
            effr[i]=sumieff[k,i]/sumitot[k,i];
            effrate=effr[i];

         calc1[i]= (1-cdf("beta", target+trange, sum(sumi[k,i],1),
sum(sumitot[k,i],-sumi[k,i])+1))/(1-target-trange);
            calc2[i]=  (cdf("beta", target+trange, sum(sumi[k,i],1),
sum(sumitot[k,i],-sumi[k,i])+1) - cdf("beta", target-trange,
sum(sumi[k,i],1), sum(sumitot[k,i],-sumi[k,i])+1))/(trange+trange);
            calc3[i]= (cdf("beta", target-trange, sum(sumi[k,i],1),
sum(sumitot[k,i],-sumi[k,i])+1)/(target-trange));
            calc4[i]= 1-cdf("beta", target, sum(sumi[k,i],1),
sum(sumitot[k,i],-sumi[k,i])+1);

            peoplec=peoplec+1;
            peopledosel[i]=sum(peopledosel[i],0.001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/1000);
      end;
    end;

      else if calc1[i]>calc2[i] and calc1[i]>calc3[i] then do;
    i=i-1;
      if i > 0 then do;
      range=sum(sumitot[k,i],0);
      do j=range+1 to range+cohortsize;
            if ranuni(1)<dr1[i] then a[i,j]=1; else a[i,j]=0;
/*a[i,j]=rand('Bernoulli', dr1[i]);*/
            sumi[k,i]=sum(sumi[k,i],a[i,j]);
            sumitot[k,i]=sum(sumitot[k,i],1);
            sumsum=sumitot[k,i];

            p1=dr1[i];
         p2=dreff[i];
            q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
           q2=(p2-q1*p1)/(1-p1);
         if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
```

```
        else aeff[i,j]=rand("binomial",q2,1);
            sumieff[k,i]=sum(sumieff[k,i],aeff[i,j]);

            dltr[i]=sumi[k,i]/sumitot[k,i];
            dltrate=dltr[i];
            effr[i]=sumieff[k,i]/sumitot[k,i];
            effrate=effr[i];

            calc1[i]= (1-cdf("beta", target+trange, sum(sumi[k,i],1),
sum(sumitot[k,i],-sumi[k,i])+1))/(1-target-trange);
            calc2[i]=  (cdf("beta", target+trange, sum(sumi[k,i],1),
sum(sumitot[k,i],-sumi[k,i])+1) - cdf("beta", target-trange,
sum(sumi[k,i],1), sum(sumitot[k,i],-sumi[k,i])+1))/(trange+trange);
            calc3[i]= (cdf("beta", target-trange, sum(sumi[k,i],1),
sum(sumitot[k,i],-sumi[k,i])+1)/(target-trange));
            calc4[i]= 1-cdf("beta", target, sum(sumi[k,i],1),
sum(sumitot[k,i],-sumi[k,i])+1);

            peoplec=peoplec+1;
            peopledosel[i]=sum(peopledosel[i],0.001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/1000);
      end;
    end;
      end;

end;

if i eq 0 then dosel[k]=20;
if i ne 0 then dosel[k]=i;

totalpop[k]=peoplec;

end;

maxdose= max(of dosel1-dosel1000);
medianpeop= median(of totalpop1-totalpop1000);
run;

/* print median number of patients on trial, avg DLT rate and
response rate in each dose level, average number of patients and
DLTs in each dose level */
proc print;
var medianpeop  dltr1-dltr10 effr1-effr10 peopledosel1-peopledosel10
dltdosel1-dltdosel10;
run;

data simi3;
set simi2 (keep=dosel1-dosel1000);

array v_[*] dosel1-dosel1000;
array counts[20];
```

```
call missing (of counts[*]);
do i = 1 to dim(v_);
     counts[v_[i]] + 1;
end;

run;

/** number of times out of 1000 that each dose level is selected for
safety before isotonic regression **/

proc print;
var counts1-counts20;
run;

/* isotonic regression for observed DLT rates */


%macro doit(k1);

%do jj=1  %to &k1;

%global max_els2;

%let onedim=Yes; *Yes/No;


%let num1=%eval((&jj-1)*10+1);
%let num2=%eval((&jj-1)*10+2);
%let num3=%eval((&jj-1)*10+3);
%let num4=%eval((&jj-1)*10+4);
%let num5=%eval((&jj-1)*10+5);
%let num6=%eval((&jj-1)*10+6);
%let num7=%eval((&jj-1)*10+7);
%let num8=%eval((&jj-1)*10+8);
%let num9=%eval((&jj-1)*10+9);
%let num10=%eval((&jj-1)*10+10);


proc iml;
use simi2 ;

read all var{sumi&num1 sumi&num2 sumi&num3 sumi&num4 sumi&num5
sumi&num6 sumi&num7 sumi&num8 sumi&num9 sumi&num10} into x;

read all var{sumitot&num1 sumitot&num2 sumitot&num3 sumitot&num4
sumitot&num5 sumitot&num6 sumitot&num7 sumitot&num8 sumitot&num9
sumitot&num10} into w;

mi=missing(x);
rowmi = mi[,+];
rowmi1=10-rowmi;
```

```
call symput('rowmi2',char(rowmi1));

/*
print mi rowmi rowmi1;
*/

create one1 from w;
append from w;
create one2 from x;
append from x;
quit;
run;

%let max_els2=&rowmi2; *Number rows/columns;

%macro assign(v,out);

data &out;
set &out;
array cols[*] col1-col&max_els2;
array s[*] &v.1-&v.&max_els2;
do i=1 to &max_els2;
%if &v=w %then %do; s[i]=max(0.0001,cols[i]); %end;
%else %do; s[i]=cols[i]; %end;
end;
keep &v:;
run;

%mend;
%assign(v=w,out=one1);
%assign(v=x,out=one2);

data one;
merge one1 one2;
array ys[*] y1-y&max_els2;
array ws[*] w1-w&max_els2;
array xs[*] x1-x&max_els2;
array zs[*] z1-z&max_els2;
array ss[*] s1-s&max_els2;
do i=1 to &max_els2;
if ws[i]>0 then ys[i]=xs[i]/ws[i];
else ys[i]=0;
zs[i]=ys[i];
ss[i]=0;
end;
drop i x:;
run;

proc datasets;
delete one1 one2;
```

```
run;

/*
proc print n width=min;
title 'orig';
var w1-w&max_els2 y1-y&max_els2;
run;
*/

%let flag=0;
%let n=0;
%do %until(&flag=1 or (&onedim=Yes and &n=1));
%let n=%eval(&n+1);

%macro rename(v,out);

data &out;
retain &v.1-&v.&max_els2;
keep &v.1-&v.&max_els2;
set one;
run;

%mend;
%rename(v=y,out=y);
%rename(v=w,out=w);
%rename(v=z,out=z);
%rename(v=s,out=s);

%if &onedim^=Yes %then %do;
%if &n^=1 %then %do;
proc iml;
use y; read all into one1;
use w; read all into one2;
use z; read all into one3;
use s; read all into one4;
new=T(one1);
new2=T(one2);
new3=T(one3);
new4=T(one4);
create one1 from new;
append from new;
create one2 from new2;
append from new2;
create one3 from new3;
append from new3;
create one4 from new4;
append from new4;
quit;
run;

%assign(v=y,out=one1);
```

```
%assign(v=w,out=one2);
%assign(v=z,out=one3);
%assign(v=s,out=one4);

data one;
merge one1 one2 one3 one4;
run;

proc datasets;
delete one1 one2 one3 one4 y w z s;
run;
%end;
%end;

data one;
set one;

%macro pav(max_els=,array=,weights=);

array _yo_[&max_els] yo1-yo&max_els;
array _w_[&max_els] w1-w&max_els;
array _y_[&max_els] y1-y&max_els;
array _z_[&max_els] z1-z&max_els;
array _s_[&max_els] s1-s&max_els;

do i=1 to &max_els;
_yo_[i]=_y_[i];
end;

%if &n=2 %then %do;
do i=1 to &max_els;
_y_[i]=_z_[i];
end;
%end;
%else %if &n>=3 %then %do;
do i=1 to &max_els;
*_y_[i]=max(0,_z_[i]+_s_[i]);
_y_[i]=_z_[i]+_s_[i];
end;
%end;

%global index ;
%if %quote(&index)=
%then
%let index=1 ;
%else
   %let index=%eval(&index+1) ;
%let pooled = _pool&index._ ;
%let parray = _parr&index._ ;
%let pwghts = _pwgt&index._ ;
```

```
array &pooled {&max_els} _TEMPORARY_ ;
array &parray {&max_els} _TEMPORARY_ ;
array &pwghts {&max_els} _TEMPORARY_ ;

if dim(&array) = 1 then Go to epav&index ;

do _pav_j_ = 1 to dim(&array) ;
&pooled(_pav_j_) = 0 ;
&parray(_pav_j_) = 0 ;
&pwghts(_pav_j_) = 0 ;
end ;

&parray(1) = &array(1) ;
&pwghts(1) = &weights(1) ;
_pav_j_ = 1 ;

do _pav_i_ = 2 to dim(&array) ;

  /* if ajacent violated, then pool */
  if (&parray(_pav_j_) > &array(_pav_i_)) then do ;
     _plwght_ = &pwghts(_pav_j_) + &weights(_pav_i_) ;
    _plval_ = ((&parray(_pav_j_)*&pwghts(_pav_j_)) +
              (&array(_pav_i_)*&weights(_pav_i_))) / _plwght_ ;

    &pooled(_pav_i_) = 1 ;
    flag=0; *added;
    if _pav_j_ > 1 then do ;
       _pav_j_ = _pav_j_ - 1 ;
       _pav_jj_ = _pav_i_ ;
       do while (flag=0 & ( &parray(_pav_j_)>_plval_) & (_pav_i_ >=
1)); *added flag=0;
       put &parray(_pav_j_)= _pav_j_= _plval_=;
         _tplval_ = _plval_ ;
         _tplwgt_ = _plwght_ ; /* tplwgt misspelled */
          do until (not &pooled(_pav_jj_)) ;
              _pav_jj_ = _pav_jj_ - 1 ;
           put _pav_jj_= &pooled(_pav_jj_)=;
              end ; /* do until */

          _plwght_ = &pwghts(_pav_j_) + _tplwgt_ ;
          _plval_  = ((&parray(_pav_j_) * &pwghts(_pav_j_)) +
                      (_tplval_*_tplwgt_)) / _plwght_ ;
         &pooled(_pav_jj_) = 1 ;
         _pav_j_ = _pav_j_ -1 ;
         if _pav_j_=0 then do; flag=1; _pav_j_=1; end; *added;
         end ; /* do while */
      _pav_j_ = _pav_j_ + 1 ;
      end ; /* if _pav_j_ > 1 */
      if flag=1 then do; *added;
         _pav_j_=_pav_j_-1; *added;
         do i=1 to _pav_i_;  *added*;
```

```
          &parray(i) = _plval_ ; *added;
          &pwghts(i) = _plwght_ ; *added;
          end;  *added*;
          end;  *added*;
      else do; *added;
      &parray(_pav_j_) = _plval_ ;
      &pwghts(_pav_j_) = _plwght_ ;
      end; *added;
      end ; /* if (&parray... */

      else do ;
          _pav_j_ = _pav_j_ + 1 ;
           &parray(_pav_j_) = &array(_pav_i_) ;
           &pwghts(_pav_j_) = &weights(_pav_i_) ;
           end ;
      end ; /* _pav_i_=2 to dim(array) */

  &array(1) = &parray(1) ;
 _pav_j_=1 ;
 _pav_jj_=1 ; /*left off underscore*/

 /* put the pooled data back into the original array */
 do _pav_j_ = 2 to dim(&array) ;
   if ^&pooled(_pav_j_) then _pav_jj_ = _pav_jj_ + 1 ;
   &array(_pav_j_) = &parray(_pav_jj_) ;
 end ;

 Epav&index:
 drop _pav_j_ _pav_i_ _pav_jj_ _plval_ _plwght_ _tplval_ _tplwgt_ ;

          ok=0;
          do i=1 to &max_els;
          %if &n=1 %then %do;
          _s_[i]=_y_[i]-_z_[i]; %end;
          %else %if &n>=2 %then %do;
          _s_[i]=_y_[i]-_z_[i]-_s_[i]; %end;
          if abs(_yo_[i]-_y_[i])>.0001 then ok=ok+1;
          end;
              run;

              %mend;

%pav(max_els=&max_els2,array=_y_,weights=_w_);

data one;
set one;
index2=1;
run;

data one;
set one end=lastrec;
```

```
by index2;
retain ok2;
if first.index2 then ok2=ok;
else ok2=ok2+ok;
if lastrec and ((mod(&n,2)^=0 and ok2=0) or &n=199) then call
symput('flag','1');
*if lastrec and (&n=51) then call symput('flag','1');

/** highest dose level with DLT rate <=0.33 after isotonic
regression **/
if y1>0.33 then mtdiso=0;
if .<y1<=0.33 then mtdiso=1;
if .<y2<=0.33 then mtdiso=2;
if .<y3<=0.33 then mtdiso=3;
if .<y4<=0.33 then mtdiso=4;
if .<y5<=0.33 then mtdiso=5;
if .<y6<=0.33 then mtdiso=6;
if .<y7<=0.33 then mtdiso=7;
if .<y8<=0.33 then mtdiso=8;
if .<y9<=0.33 then mtdiso=9;
if .<y10<=0.33 then mtdiso=10;

drop i index2 ok:;
run;

/*
proc print; var  w1-w&max_els2 y1-y&max_els2; title "Iter &n After
Iso"; run;
*/

data finalmtdiso&jj;
set one;
keep mtdiso;
run;

%end;

%end;
%mend;


%doit(1000);

data finaliso;
set finalmtdiso1 - finalmtdiso1000;
run;

/** number of times out of 1000 that each dose level is selected for
safety after isotonic regression (highest dose level at which the
smoothed DLT rate is <=0.33) **/
```

```
proc freq;
tables mtdiso;
run;

***********;
/* isotonic regression for observed response rates */

%macro doiteff(k1);

%do jj=1  %to &k1;

%global max_els2;

%let onedim=Yes; *Yes/No;


%let num1=%eval((&jj-1)*10+1);
%let num2=%eval((&jj-1)*10+2);
%let num3=%eval((&jj-1)*10+3);
%let num4=%eval((&jj-1)*10+4);
%let num5=%eval((&jj-1)*10+5);
%let num6=%eval((&jj-1)*10+6);
%let num7=%eval((&jj-1)*10+7);
%let num8=%eval((&jj-1)*10+8);
%let num9=%eval((&jj-1)*10+9);
%let num10=%eval((&jj-1)*10+10);


proc iml;
use simi2 ;

read all var{sumieff&num1 sumieff&num2 sumieff&num3 sumieff&num4
sumieff&num5 sumieff&num6 sumieff&num7 sumieff&num8 sumieff&num9
sumieff&num10} into x;

read all var{sumitot&num1 sumitot&num2 sumitot&num3 sumitot&num4
sumitot&num5 sumitot&num6 sumitot&num7 sumitot&num8 sumitot&num9
sumitot&num10} into w;

mi=missing(x);
rowmi = mi[,+];
rowmi1=10-rowmi;

call symput('rowmi2',char(rowmi1));

/*
print mi rowmi rowmi1;
*/

create one1 from w;
append from w;
```

```
create one2 from x;
append from x;
quit;
run;

%let max_els2=&rowmi2; *Number rows/columns;

%macro assign(v,out);

data &out;
set &out;
array cols[*] col1-col&max_els2;
array s[*] &v.1-&v.&max_els2;
do i=1 to &max_els2;
%if &v=w %then %do; s[i]=max(0.0001,cols[i]); %end;
%else %do; s[i]=cols[i]; %end;
end;
keep &v:;
run;

%mend;
%assign(v=w,out=one1);
%assign(v=x,out=one2);

data one;
merge one1 one2;
array ys[*] y1-y&max_els2;
array ws[*] w1-w&max_els2;
array xs[*] x1-x&max_els2;
array zs[*] z1-z&max_els2;
array ss[*] s1-s&max_els2;
do i=1 to &max_els2;
if ws[i]>0 then ys[i]=xs[i]/ws[i];
else ys[i]=0;
zs[i]=ys[i];
ss[i]=0;
end;
drop i x:;
run;

proc datasets;
delete one1 one2;
run;

/*
proc print n width=min;
title 'orig';
var w1-w&max_els2 y1-y&max_els2;
run;
*/
```

```
%let flag=0;
%let n=0;
%do %until(&flag=1 or (&onedim=Yes and &n=1));
%let n=%eval(&n+1);

%macro rename(v,out);

data &out;
retain &v.1-&v.&max_els2;
keep &v.1-&v.&max_els2;
set one;
run;

%mend;
%rename(v=y,out=y);
%rename(v=w,out=w);
%rename(v=z,out=z);
%rename(v=s,out=s);

%if &onedim^=Yes %then %do;
%if &n^=1 %then %do;
proc iml;
use y; read all into one1;
use w; read all into one2;
use z; read all into one3;
use s; read all into one4;
new=T(one1);
new2=T(one2);
new3=T(one3);
new4=T(one4);
create one1 from new;
append from new;
create one2 from new2;
append from new2;
create one3 from new3;
append from new3;
create one4 from new4;
append from new4;
quit;
run;

%assign(v=y,out=one1);
%assign(v=w,out=one2);
%assign(v=z,out=one3);
%assign(v=s,out=one4);

data one;
merge one1 one2 one3 one4;
run;

proc datasets;
```

```
delete one1 one2 one3 one4 y w z s;
run;
%end;
%end;

data one;
set one;

%macro pav(max_els=,array=,weights=);

array _yo_[&max_els] yo1-yo&max_els;
array _w_[&max_els] w1-w&max_els;
array _y_[&max_els] y1-y&max_els;
array _z_[&max_els] z1-z&max_els;
array _s_[&max_els] s1-s&max_els;

do i=1 to &max_els;
_yo_[i]=_y_[i];
end;

%if &n=2 %then %do;
do i=1 to &max_els;
_y_[i]=_z_[i];
end;
%end;
%else %if &n>=3 %then %do;
do i=1 to &max_els;
*_y_[i]=max(0,_z_[i]+_s_[i]);
_y_[i]=_z_[i]+_s_[i];
end;
%end;

%global index ;
%if %quote(&index)=
%then
%let index=1 ;
%else
   %let index=%eval(&index+1) ;
%let pooled = _pool&index._ ;
%let parray = _parr&index._ ;
%let pwghts = _pwgt&index._ ;

array &pooled {&max_els} _TEMPORARY_ ;
array &parray {&max_els} _TEMPORARY_ ;
array &pwghts {&max_els} _TEMPORARY_ ;

if dim(&array) = 1 then Go to epav&index ;

do _pav_j_ = 1 to dim(&array) ;
&pooled(_pav_j_) = 0 ;
&parray(_pav_j_) = 0 ;
```

```
&pwghts(_pav_j_) = 0 ;
end ;

&parray(1) = &array(1) ;
&pwghts(1) = &weights(1) ;
_pav_j_ = 1 ;

do _pav_i_ = 2 to dim(&array) ;

  /* if ajacent violated, then pool */
  if (&parray(_pav_j_) > &array(_pav_i_)) then do ;
     _plwght_ = &pwghts(_pav_j_) + &weights(_pav_i_) ;
    _plval_ = ((&parray(_pav_j_)*&pwghts(_pav_j_)) +
              (&array(_pav_i_)*&weights(_pav_i_))) / _plwght_ ;

    &pooled(_pav_i_) = 1 ;
    flag=0; *added;
    if _pav_j_ > 1 then do ;
       _pav_j_ = _pav_j_ - 1 ;
       _pav_jj_ = _pav_i_ ;
       do while (flag=0 & ( &parray(_pav_j_)>_plval_) & (_pav_i_ >=
1)); *added flag=0;
       put &parray(_pav_j_)= _pav_j_= _plval_=;
         _tplval_ = _plval_ ;
         _tplwgt_ = _plwght_ ; /* tplwgt misspelled */
          do until (not &pooled(_pav_jj_)) ;
              _pav_jj_ = _pav_jj_ - 1 ;
           put _pav_jj_= &pooled(_pav_jj_)=;
              end ; /* do until */

          _plwght_ = &pwghts(_pav_j_) + _tplwgt_ ;
          _plval_  = ((&parray(_pav_j_) * &pwghts(_pav_j_)) +
                      (_tplval_*_tplwgt_)) / _plwght_ ;
         &pooled(_pav_jj_) = 1 ;
         _pav_j_ = _pav_j_ -1 ;
         if _pav_j_=0 then do; flag=1; _pav_j_=1; end; *added;
         end ; /* do while */
      _pav_j_ = _pav_j_ + 1 ;
      end ; /* if _pav_j_ > 1 */
      if flag=1 then do; *added;
         _pav_j_=_pav_j_-1; *added;
         do i=1 to _pav_i_;  *added*;
         &parray(i) = _plval_ ; *added;
         &pwghts(i) = _plwght_ ; *added;
         end;  *added*;
         end;  *added*;
     else do; *added;
     &parray(_pav_j_) = _plval_ ;
     &pwghts(_pav_j_) = _plwght_ ;
     end; *added;
     end ; /* if (&parray... */
```

```
    else do ;
        _pav_j_ = _pav_j_ + 1 ;
         &parray(_pav_j_) = &array(_pav_i_) ;
         &pwghts(_pav_j_) = &weights(_pav_i_) ;
         end ;
    end ; /* _pav_i_=2 to dim(array) */

 &array(1) = &parray(1) ;
_pav_j_=1 ;
_pav_jj_=1 ; /*left off underscore*/

/* put the pooled data back into the original array */
do _pav_j_ = 2 to dim(&array) ;
  if ^&pooled(_pav_j_) then _pav_jj_ = _pav_jj_ + 1 ;
  &array(_pav_j_) = &parray(_pav_jj_) ;
end ;

Epav&index:
drop _pav_j_ _pav_i_ _pav_jj_ _plval_ _plwght_ _tplval_ _tplwgt_ ;

          ok=0;
          do i=1 to &max_els;
          %if &n=1 %then %do;
          _s_[i]=_y_[i]-_z_[i]; %end;
          %else %if &n>=2 %then %do;
          _s_[i]=_y_[i]-_z_[i]-_s_[i]; %end;
          if abs(_yo_[i]-_y_[i])>.0001 then ok=ok+1;
          end;
              run;

              %mend;

%pav(max_els=&max_els2,array=_y_,weights=_w_);

data one;
set one;
index2=1;
run;

data one;
set one end=lastrec;
by index2;
retain ok2;
if first.index2 then ok2=ok;
else ok2=ok2+ok;
if lastrec and ((mod(&n,2)^=0 and ok2=0) or &n=199) then call
symput('flag','1');
*if lastrec and (&n=51) then call symput('flag','1');
```

```
/** lowest dose level with response rate >=0.4 after isotonic
regression **/
effiso=0;
if y1>=0.4 then effiso=1;
else if y2>=0.4 then effiso=2;
else if y3>=0.4 then effiso=3;
else if y4>=0.4 then effiso=4;
else if y5>=0.4 then effiso=5;
else if y6>=0.4 then effiso=6;
else if y7>=0.4 then effiso=7;
else if y8>=0.4 then effiso=8;
else if y9>=0.4 then effiso=9;
else if y10>=0.4 then effiso=10;

drop i index2 ok:;
run;

/*
proc print; var  w1-w&max_els2 y1-y&max_els2; title "Iter &n After
Iso"; run;
*/

data finaleffiso&jj;
set one;
keep effiso;
run;

%end;

%end;
%mend;


%doiteff(1000);

data finaliso1;
set finaleffiso1 - finaleffiso1000;
run;

/** number of times out of 1000 that each dose level is selected for
efficacy after isotonic regression (lowest dose level at which the
smoothed response rate is >=0.4) **/
proc freq;
tables effiso;
run;
```

SAS code for the extended TEQR design for isotonic regression on the difference in observed response rates between adjacent dose levels from each simulation (used in determining the optimal dose for efficacy and safety for an umbrella-shaped dose-response curve):

```
data simi2 aa4;
call streaminit(1);

array a{10,100};
array aeff{10,100};
array sumi{1000,10};
array sumieff{1000,10};
array sumitot{1000,10};
array dltr{10};
array effr{10};
array dosel{1000};
array totalpop{1000};
array dr1{10};
array dreff{10};
array peopledosel{10};
array dltdosel{10};

maxdlt=0.34;
target=0.2;
trange=0.05;
maxmtdss=50;
cohortsize=5;

er=0.6;
r=0;
flagytox=0;
flagyeff=0;
flagyboth=0;

do k=1 to 1000;

/** True toxicity rates and true response rates at each dose **/
dr1[1]=0.01;
dr1[2] =0.02199;
dr1[3]=0.06165;
dr1[4]=0.2;
dr1[5]=0.55412;
dr1[6]=0.88790;
dr1[7]=0.98936;
dr1[8]=0.99959;
dr1[9]=0.99999;
dr1[10]=1;

dreff[1]=0.1;
dreff[2]=0.3;
dreff[3]=0.4;
dreff[4]=0.45;
dreff[5]=0.55;
dreff[6]=0.6;
```

```
dreff[7]=0.65;


do i1=1 to 10;
do j1=1 to maxmtdss;
        a[i1,j1]=.;
        aeff[i1,j1]=.;
            sumi[k,1]=0;
            sumieff[k,1]=0;
            sumitot[k,1]=0;
            dltr[i1]=.;
            effr[i1]=.;
            dltr[1]=0;
            effr[1]=0;
end;
end;


i=2; /** start at dose level 2 to allow de-escalation **/
sumi[k,2]=0;
sumieff[k,2]=0;
sumitot[k,2]=0;
peoplec=0;

do j=1 to cohortsize;
            a[i,j]=rand('Bernoulli', dr1[i]);
            sumi[k,i]=sumi[k,i]+a[i,j];
            sumitot[k,i]=sumitot[k,i]+1;

            p1=dr1[i];
        p2=dreff[i];
            q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
          q2=(p2-q1*p1)/(1-p1);
        if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
        else aeff[i,j]=rand("binomial",q2,1);
            sumieff[k,i]=sumieff[k,i]+aeff[i,j];

            dltr[i]=sumi[k,i]/sumitot[k,i];
            effr[i]=sumieff[k,i]/sumitot[k,i];

            peoplec=peoplec+1;
            peopledosel[i]=sum(peopledosel[i],0.001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/1000);
end;

/** dosing decisions of the TEQR design (escalate, stay at the same
dose, de-escalate) performed until the maximum MTD sample size is
reached or we need to de-escalate beyond dose level 1 **/

 do until ((sumsum=maxmtdss and dltrate<maxdlt) or i=0 /* or
effrate>=er*/);
```

```
  if 0<=dltr[i]<(target-trange) and dltr[i+1]<maxdlt then do;
  i=i+1;
    range=sum(sumitot[k,i],0);
    do j=range+1 to range+cohortsize;
          a[i,j]=rand('Bernoulli', dr1[i]);
          sumi[k,i]=sum(sumi[k,i],a[i,j]);
          sumitot[k,i]=sum(sumitot[k,i],1);
          sumsum=sumitot[k,i];

          p1=dr1[i];
      p2=dreff[i];
          q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
        q2=(p2-q1*p1)/(1-p1);
      if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
      else aeff[i,j]=rand("binomial",q2,1);
          sumieff[k,i]=sum(sumieff[k,i],aeff[i,j]);

          dltr[i]=sumi[k,i]/sumitot[k,i];
          dltrate=dltr[i];
          effr[i]=sumieff[k,i]/sumitot[k,i];
          effrate=effr[i];

          peoplec=peoplec+1;
          peopledosel[i]=sum(peopledosel[i],0.001);
          dltdosel[i]=sum(dltdosel[i],a[i,j]/1000);
    end;
  end;
else if 0<=dltr[i]<(target-trange) and dltr[i+1]>=maxdlt then do;
    range=sum(sumitot[k,i],0);
    do j=range+1 to range+cohortsize;
          a[i,j]=rand('Bernoulli', dr1[i]);
          sumi[k,i]=sum(sumi[k,i],a[i,j]);
          sumitot[k,i]=sum(sumitot[k,i],1);
          sumsum=sumitot[k,i];

          p1=dr1[i];
      p2=dreff[i];
          q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
        q2=(p2-q1*p1)/(1-p1);
      if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
      else aeff[i,j]=rand("binomial",q2,1);
          sumieff[k,i]=sum(sumieff[k,i],aeff[i,j]);

          dltr[i]=sumi[k,i]/sumitot[k,i];
          dltrate=dltr[i];
          effr[i]=sumieff[k,i]/sumitot[k,i];
          effrate=effr[i];

          peoplec=peoplec+1;
          peopledosel[i]=sum(peopledosel[i],0.001);
          dltdosel[i]=sum(dltdosel[i],a[i,j]/1000);
    end;
  end;
```

```
else if (target-trange)<=dltr[i]<=(target+trange) then do;
range=sum(sumitot[k,i],0);
     do j=range+1 to range+cohortsize;
           a[i,j]=rand('Bernoulli', dr1[i]);
           sumi[k,i]=sum(sumi[k,i],a[i,j]);
           sumitot[k,i]=sum(sumitot[k,i],1);
           sumsum=sumitot[k,i];

           p1=dr1[i];
       p2=dreff[i];
           q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
         q2=(p2-q1*p1)/(1-p1);
       if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
       else aeff[i,j]=rand("binomial",q2,1);
           sumieff[k,i]=sum(sumieff[k,i],aeff[i,j]);

           dltr[i]=sumi[k,i]/sumitot[k,i];
           dltrate=dltr[i];
           effr[i]=sumieff[k,i]/sumitot[k,i];
           effrate=effr[i];

           peoplec=peoplec+1;
           peopledosel[i]=sum(peopledosel[i],0.001);
           dltdosel[i]=sum(dltdosel[i],a[i,j]/1000);
     end;
   end;

     else if (target+trange)<dltr[i]<=1 then do;
   i=i-1;
     if i > 0 then do;
     range=sum(sumitot[k,i],0);
     do j=range+1 to range+cohortsize;
           a[i,j]=rand('Bernoulli', dr1[i]);
           sumi[k,i]=sum(sumi[k,i],a[i,j]);
           sumitot[k,i]=sum(sumitot[k,i],1);
           sumsum=sumitot[k,i];

           p1=dr1[i];
       p2=dreff[i];
           q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
         q2=(p2-q1*p1)/(1-p1);
       if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
       else aeff[i,j]=rand("binomial",q2,1);
           sumieff[k,i]=sum(sumieff[k,i],aeff[i,j]);

           dltr[i]=sumi[k,i]/sumitot[k,i];
           dltrate=dltr[i];
           effr[i]=sumieff[k,i]/sumitot[k,i];
           effrate=effr[i];

           peoplec=peoplec+1;
           peopledosel[i]=sum(peopledosel[i],0.001);
           dltdosel[i]=sum(dltdosel[i],a[i,j]/1000);
     end;
```

```
    end;
      end;

end;

if i ne 0 then dosel[k]=i;
if i eq 0 then dosel[k]=20;

totalpop[k]=peoplec;
if k>=1 then output aa4;
end;

maxdose= max(of dosel1-dosel1000);
medianpeop= median(of totalpop1-totalpop1000);
run;


data aa5;
set aa4;
keep effr1-effr10;
run;

data aa6;
set aa5;
if effr1=effr2 then effr2=effr2+0.001;
if effr1=effr3 then effr3=effr3+0.001;
if effr1=effr4 then effr4=effr4+0.001;
if effr1=effr5 then effr5=effr5+0.001;
if effr1=effr6 then effr6=effr6+0.001;
if effr2=effr3 then effr3=effr3+0.001;
if effr2=effr4 then effr4=effr4+0.001;
if effr2=effr5 then effr5=effr5+0.001;
if effr2=effr6 then effr6=effr6+0.001;
if effr3=effr4 then effr4=effr4+0.001;
if effr3=effr5 then effr5=effr5+0.001;
if effr3=effr6 then effr6=effr6+0.001;
if effr4=effr5 then effr5=effr5+0.001;
if effr4=effr6 then effr6=effr6+0.001;
if effr5=effr6 then effr6=effr6+0.001;

run;

/** dataset with difference in observed response rates between adjacent
dose levels **/
data aa7;
set aa5;
effra1=effr1-effr2;
effra2=effr2-effr3;
effra3=effr3-effr4;
effra4=effr4-effr5;
effra5=effr5-effr6;
effra6=effr6-effr7;
effra7=effr7-effr8;
effra8=effr8-effr9;
```

```
effra9=effr9-effr10;
run;

data aa9;
set aa7;
if effra1 ne . then weight1=1;
if effra2 ne . then weight2=1;
if effra3 ne . then weight3=1;
if effra4 ne . then weight4=1;
if effra5 ne . then weight5=1;
if effra6 ne . then weight6=1;
if effra7 ne . then weight7=1;
if effra8 ne . then weight8=1;
if effra9 ne . then weight9=1;
run;


/** Data sets with difference in observed response rates between
adjacent dose levels from each simulation to be used in the isotonic
regression macro **/
%macro multids;

  %do I = 1 %to 1000;

     data Data&I;

       set aa9 (keep=effra1-effra9 weight1-weight9);

       if _N_=&I then output;

     run;

     %end;

%mend;



%multids;

/** isotonic regression on the differences in observed response rates
between adjacent dose levels **/

%macro doiteff(k1);

%do jj=1  %to &k1;

%global max_els2;

%let onedim=Yes; *Yes/No;

proc iml;
use data&jj ;
```

```
read all var{effra1 effra2 effra3 effra4 effra5 effra6 effra7 effra8
effra9} into x;

read all var{weight1 weight2 weight3 weight4 weight5 weight6 weight7
weight8 weight9} into w;

mi=missing(x);
rowmi = mi[,+];
rowmi1=9-rowmi;

call symput('rowmi2',char(rowmi1));

/*
print mi rowmi rowmi1;
*/

create one1 from w;
append from w;
create one2 from x;
append from x;
quit;
run;

%let max_els2=&rowmi2; *Number rows/columns;

%macro assign(v,out);

data &out;
set &out;
array cols[*] col1-col&max_els2;
array s[*] &v.1-&v.&max_els2;
do i=1 to &max_els2;
%if &v=w %then %do; s[i]=max(0.0001,cols[i]); %end;
%else %do; s[i]=cols[i]; %end;
end;
keep &v:;
run;

%mend;
%assign(v=w,out=one1);
%assign(v=x,out=one2);

data one;
merge one1 one2;
array ys[*] y1-y&max_els2;
array ws[*] w1-w&max_els2;
array xs[*] x1-x&max_els2;
array zs[*] z1-z&max_els2;
array ss[*] s1-s&max_els2;
do i=1 to &max_els2;
if ws[i]>0 then ys[i]=xs[i]/ws[i];
else ys[i]=0;
zs[i]=ys[i];
ss[i]=0;
```

```
end;
drop i x:;
run;

proc datasets;
delete one1 one2;
run;

/*
proc print n width=min;
title 'orig';
var w1-w&max_els2 y1-y&max_els2;
run;
*/

%let flag=0;
%let n=0;
%do %until(&flag=1 or (&onedim=Yes and &n=1));
%let n=%eval(&n+1);

%macro rename(v,out);

data &out;
retain &v.1-&v.&max_els2;
keep &v.1-&v.&max_els2;
set one;
run;

%mend;
%rename(v=y,out=y);
%rename(v=w,out=w);
%rename(v=z,out=z);
%rename(v=s,out=s);

%if &onedim^=Yes %then %do;
%if &n^=1 %then %do;
proc iml;
use y; read all into one1;
use w; read all into one2;
use z; read all into one3;
use s; read all into one4;
new=T(one1);
new2=T(one2);
new3=T(one3);
new4=T(one4);
create one1 from new;
append from new;
create one2 from new2;
append from new2;
create one3 from new3;
append from new3;
create one4 from new4;
append from new4;
quit;
```

```
run;

%assign(v=y,out=one1);
%assign(v=w,out=one2);
%assign(v=z,out=one3);
%assign(v=s,out=one4);

data one;
merge one1 one2 one3 one4;
run;

proc datasets;
delete one1 one2 one3 one4 y w z s;
run;
%end;
%end;

data one;
set one;

%macro pav(max_els=,array=,weights=);

array _yo_[&max_els] yo1-yo&max_els;
array _w_[&max_els] w1-w&max_els;
array _y_[&max_els] y1-y&max_els;
array _z_[&max_els] z1-z&max_els;
array _s_[&max_els] s1-s&max_els;

do i=1 to &max_els;
_yo_[i]=_y_[i];
end;

%if &n=2 %then %do;
do i=1 to &max_els;
_y_[i]=_z_[i];
end;
%end;
%else %if &n>=3 %then %do;
do i=1 to &max_els;
*_y_[i]=max(0,_z_[i]+_s_[i]);
_y_[i]=_z_[i]+_s_[i];
end;
%end;

%global index ;
%if %quote(&index)=
%then
%let index=1 ;
%else
   %let index=%eval(&index+1) ;
%let pooled = _pool&index._ ;
%let parray = _parr&index._ ;
%let pwghts = _pwgt&index._ ;
```

```
array &pooled {&max_els} _TEMPORARY_ ;
array &parray {&max_els} _TEMPORARY_ ;
array &pwghts {&max_els} _TEMPORARY_ ;

if dim(&array) = 1 then Go to epav&index ;

do _pav_j_ = 1 to dim(&array) ;
&pooled(_pav_j_) = 0 ;
&parray(_pav_j_) = 0 ;
&pwghts(_pav_j_) = 0 ;
end ;

&parray(1) = &array(1) ;
&pwghts(1) = &weights(1) ;
_pav_j_ = 1 ;

do _pav_i_ = 2 to dim(&array) ;

  /* if ajacent violated, then pool */
  if (&parray(_pav_j_) > &array(_pav_i_)) then do ;
     _plwght_ = &pwghts(_pav_j_) + &weights(_pav_i_) ;
    _plval_ = ((&parray(_pav_j_)*&pwghts(_pav_j_)) +
              (&array(_pav_i_)*&weights(_pav_i_))) / _plwght_ ;

    &pooled(_pav_i_) = 1 ;
    flag=0; *added;
    if _pav_j_ > 1 then do ;
       _pav_j_ = _pav_j_ - 1 ;
       _pav_jj_ = _pav_i_ ;
       do while (flag=0 & ( &parray(_pav_j_)>_plval_) & (_pav_i_ >=
1)); *added flag=0;
       put &parray(_pav_j_)= _pav_j_= _plval_=;
         _tplval_ = _plval_ ;
         _tplwgt_ = _plwght_ ; /* tplwgt misspelled */
          do until (not &pooled(_pav_jj_)) ;
              _pav_jj_ = _pav_jj_ - 1 ;
           put _pav_jj_= &pooled(_pav_jj_)=;
              end ; /* do until */

          _plwght_ = &pwghts(_pav_j_) + _tplwgt_ ;
          _plval_  = ((&parray(_pav_j_) * &pwghts(_pav_j_)) +
                      (_tplval_*_tplwgt_)) / _plwght_ ;
         &pooled(_pav_jj_) = 1 ;
         _pav_j_ = _pav_j_ -1 ;
         if _pav_j_=0 then do; flag=1; _pav_j_=1; end; *added;
         end ; /* do while */
      _pav_j_ = _pav_j_ + 1 ;
      end ; /* if _pav_j_ > 1 */
      if flag=1 then do; *added;
         _pav_j_=_pav_j_-1; *added;
         do i=1 to _pav_i_;  *added*;
         &parray(i) = _plval_ ; *added;
         &pwghts(i) = _plwght_ ; *added;
         end;  *added*;
```

```
        end;  *added*;
    else do; *added;
    &parray(_pav_j_) = _plval_ ;
    &pwghts(_pav_j_) = _plwght_ ;
    end; *added;
    end ; /* if (&parray... */

    else do ;
        _pav_j_ = _pav_j_ + 1 ;
         &parray(_pav_j_) = &array(_pav_i_) ;
         &pwghts(_pav_j_) = &weights(_pav_i_) ;
         end ;
    end ; /* _pav_i_=2 to dim(array) */

 &array(1) = &parray(1) ;
_pav_j_=1 ;
_pav_jj_=1 ; /*left off underscore*/

/* put the pooled data back into the original array */
do _pav_j_ = 2 to dim(&array) ;
  if ^&pooled(_pav_j_) then _pav_jj_ = _pav_jj_ + 1 ;
  &array(_pav_j_) = &parray(_pav_jj_) ;
end ;

Epav&index:
drop _pav_j_ _pav_i_ _pav_jj_ _plval_ _plwght_ _tplval_ _tplwgt_ ;

          ok=0;
          do i=1 to &max_els;
          %if &n=1 %then %do;
          _s_[i]=_y_[i]-_z_[i]; %end;
          %else %if &n>=2 %then %do;
          _s_[i]=_y_[i]-_z_[i]-_s_[i]; %end;
          if abs(_yo_[i]-_y_[i])>.0001 then ok=ok+1;
          end;
              run;

              %mend;

%pav(max_els=&max_els2,array=_y_,weights=_w_);

data one;
set one;
index2=1;
run;

data one;
set one end=lastrec;
by index2;
retain ok2;
if first.index2 then ok2=ok;
else ok2=ok2+ok;
if lastrec and ((mod(&n,2)^=0 and ok2=0) or &n=199) then call
symput('flag','1');
```

```
*if lastrec and (&n=51) then call symput('flag','1');

effiso=0;
if y1>=0 then effiso=1;
if y2=0 then effiso=2;
if y3=0 then effiso=3;
if y4=0 then effiso=4;
if y5=0 then effiso=5;
if y6=0 then effiso=6;

if y1<0 and y2>0 and abs(y1)<=abs(y2) then effiso=1;
if y1<0 and y2>0 and abs(y1)>abs(y2) then effiso=2;

if y2<0 and y3>0 and abs(y2)<=abs(y3) then effiso=2;
if y2<0 and y3>0 and abs(y2)>abs(y3) then effiso=3;

if y3<0 and y4>0 and abs(y3)<=abs(y4) then effiso=3;
if y3<0 and y4>0 and abs(y3)>abs(y4) then effiso=4;

if y4<0 and y5>0 and abs(y4)<=abs(y5) then effiso=4;
if y4<0 and y5>0 and abs(y4)>abs(y5) then effiso=5;

if y5<0 and y6>0 and abs(y5)<=abs(y6) then effiso=5;
if y5<0 and y6>0 and abs(y5)>abs(y6) then effiso=6;

if .<y1<0 and y2 eq .  then effiso=1;
if .<y2<0 and y3 eq .  then effiso=2;
if .<y3<0 and y4 eq .  then effiso=3;
if .<y4<0 and y5 eq .  then effiso=4;
if .<y5<0 and y6 eq .  then effiso=5;
if .<y6<0 and y7 eq .  then effiso=6;
drop i index2 ok:;
run;

/*
proc print; var  w1-w&max_els2 y1-y&max_els2; title "Iter &n After
Iso"; run;
*/

data finaleffiso&jj;
set one;
run;

%end;

%end;
%mend;


%doiteff(1000);

data finaliso1;
set finaleffiso1 - finaleffiso1000;
if effiso=1 and y1=0 then peak=2;
```

```
if effiso=1 and y1>0 then peak=1;
if effiso=1 and y1<0 and y2 ne . then peak=2;
if effiso=1 and y1<0 and y2 eq . then peak=0;
if effiso=2 and y2=0 then peak=3;
if effiso=2 and y2>0 then peak=2;
if effiso=2 and y2<0 and y3 ne . then peak=3;
if effiso=2 and y2<0 and y3 eq . then peak=0;
if effiso=3 and y3=0 then peak=4;
if effiso=3 and y3>0 then peak=3;
if effiso=3 and y3<0 and y4 ne . then peak=4;
if effiso=3 and y3<0 and y4 eq . then peak=0;
if effiso=4 and y4=0 then peak=5;
if effiso=4 and y4>0 then peak=4;
if effiso=4 and y4<0 and y5 ne . then peak=5;
if effiso=4 and y4<0 and y5 eq . then peak=0;
if effiso=5 and y5=0 then peak=6;
if effiso=5 and y5>0 then peak=5;
if effiso=5 and y5<0 and y6 ne . then peak=6;
if effiso=5 and y5<0 and y6 eq . then peak=0;



keep y1-y9 yo1-yo9 peak effiso;
run;

proc print data=finaliso1;
var y1-y5 yo1-yo5 effiso peak;
run;


/** Number of times out of 1000 simulations that each dose level is
choosen as the peak **/
proc freq;
tables peak;
run;

/** Code for finding the optimal dose for an umbrella shaped dose-
response curve **/

data finaliso1;
set finaliso1;
retain id;
id = sum(id, 1);
run;

/** Data set with the dose level that is selected as the highest dose
that is safe after isotonic regression - can be obtained from the first
TEQR program in "Codes for Chapter 4". **/

PROC IMPORT OUT= WORK.mtdiso
            DATAFILE= "D:\Documents and
Settings\rananthakrishnan\Deskto
p\mtdiso.xls"
```

```
            DBMS=EXCEL REPLACE;
     RANGE="mtdiso";
     GETNAMES=YES;
     MIXED=NO;
     SCANTEXT=YES;
     USEDATE=YES;
     SCANTIME=YES;
RUN;

data aa5;
set aa5;
retain id;
id=sum(id, 1);
run;

data mtdiso;
set mtdiso;
retain id;
id=sum(id, 1);
run;

data finaliso2;
merge finaliso1 mtdiso aa5;
by id;
if peak=0 then optimal=0;

if peak=6 and effr6>=0.4 and mtdiso>=6 then optimal=6;
if peak=6 and effr6<0.4 and mtdiso>=6 then optimal=0;
if peak=6 and mtdiso=5 and effr5>=0.4 then optimal=5;
if peak=6 and mtdiso=5 and effr5<0.4 then optimal=0;
if peak=6 and mtdiso=4 and effr4>=0.4 then optimal=4;
if peak=6 and mtdiso=4 and effr4<0.4 then optimal=0;
if peak=6 and mtdiso=3 and effr3>=0.4 then optimal=3;
if peak=6 and mtdiso=3 and effr3<0.4 then optimal=0;
if peak=6 and mtdiso=2 and effr2>=0.4 then optimal=2;
if peak=6 and mtdiso=2 and effr2<0.4 then optimal=0;
if peak=6 and mtdiso=1 and effr1>=0.4 then optimal=1;
if peak=6 and mtdiso=1 and effr1<0.4 then optimal=0;

if peak=5 and effr5>=0.4 and mtdiso>=5 then optimal=5;
if peak=5 and effr5<0.4 and mtdiso>=5 then optimal=0;
if peak=5 and mtdiso=4 and effr4>=0.4 then optimal=4;
if peak=5 and mtdiso=4 and effr4<0.4 then optimal=0;
if peak=5 and mtdiso=3 and effr3>=0.4 then optimal=3;
if peak=5 and mtdiso=3 and effr3<0.4 then optimal=0;
if peak=5 and mtdiso=2 and effr2>=0.4 then optimal=2;
if peak=5 and mtdiso=2 and effr2<0.4 then optimal=0;
if peak=5 and mtdiso=1 and effr1>=0.4 then optimal=1;
if peak=5 and mtdiso=1 and effr1<0.4 then optimal=0;

if peak=4 and effr4>=0.4 and mtdiso>=4 then optimal=4;
if peak=4 and effr4<0.4 and mtdiso>=4 then optimal=0;
if peak=4 and mtdiso=3 and effr3>=0.4 then optimal=3;
if peak=4 and mtdiso=3 and effr3<0.4 then optimal=0;
```

```
if peak=4 and mtdiso=2 and effr2>=0.4 then optimal=2;
if peak=4 and mtdiso=2 and effr2<0.4 then optimal=0;
if peak=4 and mtdiso=1 and effr1>=0.4 then optimal=1;
if peak=4 and mtdiso=1 and effr1<0.4 then optimal=0;

if peak=3 and effr3>=0.4 and mtdiso>=3 then optimal=3;
if peak=3 and effr3<0.4 and mtdiso>=3 then optimal=0;
if peak=3 and mtdiso=2 and effr2>=0.4 then optimal=2;
if peak=3 and mtdiso=2 and effr2<0.4 then optimal=0;
if peak=3 and mtdiso=1 and effr1>=0.4 then optimal=1;
if peak=3 and mtdiso=1 and effr1<0.4 then optimal=0;

if peak=2 and effr2>=0.4 and mtdiso>=2 then optimal=2;
if peak=2 and effr2<0.4 and mtdiso>=2 then optimal=0;
if peak=2 and mtdiso=1 and effr1>=0.4 then optimal=1;
if peak=2 and mtdiso=1 and effr1<0.4 then optimal=0;

if peak=1 and effr1>=0.4 and mtdiso>=1 then optimal=1;
if peak=1 and effr1<0.4 and mtdiso>=1 then optimal=0;

run;

/** Number of times out of 1000 simulations that each dose level is
chosen as the optimal dose for safety and efficacy **/
proc freq;
tables optimal;
run;
```

SAS code for the extended TEQR design for obtaining the observed response rate at each dose level and the difference in observed response rates between adjacent dose levels from each simulation:

```
data simi2 aa4;
call streaminit(1);

array a{10,100};
array aeff{10,100};
array sumi{1000,10};
array sumieff{1000,10};
array sumitot{1000,10};
array dltr{10};
array effr{10};
array dosel{1000};
array totalpop{1000};
array dr1{10};
array dreff{10};
array peopledosel{10};
array dltdosel{10};

maxdlt=0.34;
target=0.2;
trange=0.05;
```

```
maxmtdss=50;
cohortsize=5;

er=0.6;
r=0;
flagytox=0;
flagyeff=0;
flagyboth=0;

do k=1 to 1000;

/** true DLT and response rates **/
dr1[1]=0.01;
dr1[2] =0.02199;
dr1[3]=0.06165;
dr1[4]=0.2;
dr1[5]=0.55412;
dr1[6]=0.88790;
dr1[7]=0.98936;
dr1[8]=0.99959;
dr1[9]=0.99999;
dr1[10]=1;

dreff[1]=0.1;
dreff[2]=0.35;
dreff[3]=0.5;
dreff[4]=0.3;
dreff[5]=0.2;
dreff[6]=0.05;
dreff[7]=0.01;


do i1=1 to 10;
do j1=1 to maxmtdss;
        a[i1,j1]=.;
        aeff[i1,j1]=.;
            sumi[k,1]=0;
            sumieff[k,1]=0;
            sumitot[k,1]=0;
            dltr[i1]=.;
            effr[i1]=.;
            dltr[1]=0;
            effr[1]=0;
end;
end;


i=2; /** start at dose level 2 to allow de-escalation **/
sumi[k,2]=0;
sumieff[k,2]=0;
sumitot[k,2]=0;
peoplec=0;

do j=1 to cohortsize;
```

```
            a[i,j]=rand('Bernoulli', dr1[i]);
            sumi[k,i]=sumi[k,i]+a[i,j];
            sumitot[k,i]=sumitot[k,i]+1;

            p1=dr1[i];
        p2=dreff[i];
            q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
          q2=(p2-q1*p1)/(1-p1);
        if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
        else aeff[i,j]=rand("binomial",q2,1);
            sumieff[k,i]=sumieff[k,i]+aeff[i,j];

            dltr[i]=sumi[k,i]/sumitot[k,i];
            effr[i]=sumieff[k,i]/sumitot[k,i];

            peoplec=peoplec+1;
            peopledosel[i]=sum(peopledosel[i],0.001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/1000);
end;

/** dosing decisions of the TEQR design (escalate, stay at the same
dose, de-escalate) performed until the maximum MTD sample size is
reached or we need to de-escalate beyond dose level 1 **/
 do until ((sumsum=maxmtdss and dltrate<maxdlt) or i=0 /* or
effrate>=er*/);


    if 0<=dltr[i]<(target-trange) and dltr[i+1]<maxdlt then do;
    i=i+1;
      range=sum(sumitot[k,i],0);
      do j=range+1 to range+cohortsize;
            a[i,j]=rand('Bernoulli', dr1[i]);
            sumi[k,i]=sum(sumi[k,i],a[i,j]);
            sumitot[k,i]=sum(sumitot[k,i],1);
            sumsum=sumitot[k,i];

            p1=dr1[i];
        p2=dreff[i];
            q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
          q2=(p2-q1*p1)/(1-p1);
        if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
        else aeff[i,j]=rand("binomial",q2,1);
            sumieff[k,i]=sum(sumieff[k,i],aeff[i,j]);

            dltr[i]=sumi[k,i]/sumitot[k,i];
            dltrate=dltr[i];
            effr[i]=sumieff[k,i]/sumitot[k,i];
            effrate=effr[i];

            peoplec=peoplec+1;
            peopledosel[i]=sum(peopledosel[i],0.001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/1000);
      end;
```

```
   end;
 else if 0<=dltr[i]<(target-trange) and dltr[i+1]>=maxdlt then do;
     range=sum(sumitot[k,i],0);
     do j=range+1 to range+cohortsize;
           a[i,j]=rand('Bernoulli', dr1[i]);
           sumi[k,i]=sum(sumi[k,i],a[i,j]);
           sumitot[k,i]=sum(sumitot[k,i],1);
           sumsum=sumitot[k,i];

           p1=dr1[i];
       p2=dreff[i];
           q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
         q2=(p2-q1*p1)/(1-p1);
       if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
       else aeff[i,j]=rand("binomial",q2,1);
           sumieff[k,i]=sum(sumieff[k,i],aeff[i,j]);

           dltr[i]=sumi[k,i]/sumitot[k,i];
           dltrate=dltr[i];
           effr[i]=sumieff[k,i]/sumitot[k,i];
           effrate=effr[i];

           peoplec=peoplec+1;
           peopledosel[i]=sum(peopledosel[i],0.001);
           dltdosel[i]=sum(dltdosel[i],a[i,j]/1000);
     end;
   end;

else if (target-trange)<=dltr[i]<=(target+trange) then do;
range=sum(sumitot[k,i],0);
     do j=range+1 to range+cohortsize;
           a[i,j]=rand('Bernoulli', dr1[i]);
           sumi[k,i]=sum(sumi[k,i],a[i,j]);
           sumitot[k,i]=sum(sumitot[k,i],1);
           sumsum=sumitot[k,i];

           p1=dr1[i];
       p2=dreff[i];
           q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
         q2=(p2-q1*p1)/(1-p1);
       if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
       else aeff[i,j]=rand("binomial",q2,1);
           sumieff[k,i]=sum(sumieff[k,i],aeff[i,j]);

           dltr[i]=sumi[k,i]/sumitot[k,i];
           dltrate=dltr[i];
           effr[i]=sumieff[k,i]/sumitot[k,i];
           effrate=effr[i];

           peoplec=peoplec+1;
           peopledosel[i]=sum(peopledosel[i],0.001);
           dltdosel[i]=sum(dltdosel[i],a[i,j]/1000);
     end;
   end;
```

```
      else if (target+trange)<dltr[i]<=1 then do;
    i=i-1;
      if i > 0 then do;
      range=sum(sumitot[k,i],0);
      do j=range+1 to range+cohortsize;
            a[i,j]=rand('Bernoulli', dr1[i]);
            sumi[k,i]=sum(sumi[k,i],a[i,j]);
            sumitot[k,i]=sum(sumitot[k,i],1);
            sumsum=sumitot[k,i];

            p1=dr1[i];
        p2=dreff[i];
            q1=p2+(r/p1)*sqrt(p1*(1-p1)*p2*(1-p2));
          q2=(p2-q1*p1)/(1-p1);
        if a[i,j]=1 then aeff[i,j]=rand("binomial", q1,1);
        else aeff[i,j]=rand("binomial",q2,1);
            sumieff[k,i]=sum(sumieff[k,i],aeff[i,j]);

            dltr[i]=sumi[k,i]/sumitot[k,i];
            dltrate=dltr[i];
            effr[i]=sumieff[k,i]/sumitot[k,i];
            effrate=effr[i];

            peoplec=peoplec+1;
            peopledosel[i]=sum(peopledosel[i],0.001);
            dltdosel[i]=sum(dltdosel[i],a[i,j]/1000);
      end;
    end;
      end;

end;

if i ne 0 then dosel[k]=i;
if i eq 0 then dosel[k]=20;

totalpop[k]=peoplec;
if k>=1 then output aa4;
end;

maxdose= max(of dosel1-dosel1000);
medianpeop= median(of totalpop1-totalpop1000);
run;


data aa5;
set aa4;
keep effr1-effr10;
run;

/** This data set provides the observed response rate at each dose
level from each simulation **/
data aa6;
set aa5;
```

```
if effr1=effr2 then effr2=effr2+0.001;
if effr1=effr3 then effr3=effr3+0.001;
if effr1=effr4 then effr4=effr4+0.001;
if effr1=effr5 then effr5=effr5+0.001;
if effr1=effr6 then effr6=effr6+0.001;
if effr2=effr3 then effr3=effr3+0.001;
if effr2=effr4 then effr4=effr4+0.001;
if effr2=effr5 then effr5=effr5+0.001;
if effr2=effr6 then effr6=effr6+0.001;
if effr3=effr4 then effr4=effr4+0.001;
if effr3=effr5 then effr5=effr5+0.001;
if effr3=effr6 then effr6=effr6+0.001;
if effr4=effr5 then effr5=effr5+0.001;
if effr4=effr6 then effr6=effr6+0.001;
if effr5=effr6 then effr6=effr6+0.001;

run;

data aa7;
set aa5;
effra1=effr1-effr2;
effra2=effr2-effr3;
effra3=effr3-effr4;
effra4=effr4-effr5;
effra5=effr5-effr6;
effra6=effr6-effr7;
effra7=effr7-effr8;
effra8=effr8-effr9;
effra9=effr9-effr10;
run;

/** This data set provides the observed difference in response rates
between adjacent dose levels from each simulation **/
data aa8;
set aa7;

if abs(effra1-effra2)<=0.001 then effra2=effra2+0.0001;
if abs(effra1-effra3)<=0.001 then effra3=effra3+0.0001;
if abs(effra1-effra4)<=0.001 then effra4=effra4+0.0001;
if abs(effra1-effra5)<=0.001 then effra5=effra5+0.0001;
if abs(effra1-effra6)<=0.001 then effra6=effra6+0.0001;
if abs(effra2-effra3)<=0.001 then effra3=effra3+0.0001;
if abs(effra2-effra4)<=0.001 then effra4=effra4+0.0001;
if abs(effra2-effra5)<=0.001 then effra5=effra5+0.0001;
if abs(effra2-effra6)<=0.001 then effra6=effra6+0.0001;
if abs(effra3-effra4)<=0.001 then effra4=effra4+0.0001;
if abs(effra3-effra5)<=0.001 then effra5=effra5+0.0001;
if abs(effra3-effra6)<=0.001 then effra6=effra6+0.0001;
if abs(effra4-effra5)<=0.001 then effra5=effra5+0.0001;
if abs(effra4-effra6)<=0.001 then effra6=effra6+0.0001;
if abs(effra5-effra6)<=0.001 then effra6=effra6+0.0001;

keep effra1-effra9;
run;
```

R code for the extended TEQR design for nearly isotonic regression applied to the observed response rates:

```
MyData <- read.csv(file="C:\\Users\\HP\\Desktop\\teqriso1.csv", header=TRUE, sep=",")
library(neariso)
lambda=c(0)

i <- 1
while (i<1001){
z<- which(is.na(MyData[i,]))[1]
y<- MyData[i,][!is.na(MyData[i,])]
y<- y[-z]
if(z > 2){

res0 <- neariso(y, lambda=lambda)
ii <- which(res0$beta >=0.4)[1]
MyData$effd[i]=ii
} else {

    MyData$effd[i]=1
}


i=i+1
}
table(MyData$effd)
```

BIBLIOGRAPHY

1. Gambacorti-Passerini C, Antolini L, Mahon FX, Guilhot F, Deininger M, Fava C, Nagler A, Della Casa CM, Morra E, Abruzzese E, D'Emilio A, Stagno F, le Coutre P, Hurtado-Monroy R, Santini V, Martino B, Pane F, Piccin A, Giraldo P, Assouline S, Durosinmi MA, Leeksma O, Pogliani EM, Puttini M, Jang E, Reiffers J, Valsecchi MG, Kim DW. Multicenter independent assessment of outcomes in chronic myeloid leukemia patients treated with imatinib. Journal of the National Cancer Institute. 2011; 103(7):553–561. Erratum in: Journal of the National Cancer Institute 2016; 108(9). pii: djw211.

2. Greene FL, Compton CC, Fritz AG, Shah JP, Winchester DP. AJCC cancer staging atlas. Springer; 2006.

3. O'Brien S, Gribben JG. Chronic lymphocytic leukemia. Informa Healthcare USA, Inc; 2008.

4. Cheson BD. Staging and response assessment in lymphomas: the new Lugano classification. Chinese Clinical Oncology. 2015; 4(1):5.

5. Rue M, Vilaprinyo E, Lee S, et al. Effectiveness of early detection on breast cancer mortality reduction in Catalonia (Spain). BMC Cancer 2009; 9:326.

6. Mascaux C, Peled N, Garg K, Kato Y, Wynes MW, Hirsch FR. Early detection and screening of lung cancer. Expert Review of Molecular Diagnostics. 2010; 10(6):799–815.

7. Chandy D, Maguire G, Aronow WS. Lung cancer: the importance of early intervention. Comprehensive Therapy. 2009 Spring; 35(1):18–23.

8. Dashwood RH. Early detection and prevention of colorectal cancer. Oncology Reports. 1999; 6(2):277–281 [review].

9. Sciubba JJ. Oral cancer: the importance of early diagnosis and treatment. American Journal of Clinical Dermatology. 2001; 2(4):239–251.

10. Wardle J, Robb K, Vernon S, Waller J. Screening for prevention and early diagnosis of cancer. American Psychologist. 2015; 70(2):119–133.

11. Wodarz D, Zauber AG. Cancer: Risk factors and random chances. Nature. 2015; 517(7536):563–564.

12. Young SR, Brooks KA, Edwards JG, Smith ST. Basic principles of cancer genetics. Journal of the South Carolina Medical Association. 1998; 94(7):299–305. Review.

13. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000; 100(1):57–70. Review.

14. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011; 144(5):646–674. Review.

15. Bertram JS. The molecular biology of cancer. Molecular Aspects of Medicine. 2000; 21(6):167–223. Review.

16. Kassen A, Hofmockel G. Molecular genetic and cell biology principles for the development of malignant tumors. Urologe A. 2000; 39(3):214–221. Review.

17. Harley CB. Telomerase is not an oncogene. Oncogene. 2002; 21(4):494–502.

18. Suresh S. Biomechanics and biophysics of cancer cells. Acta Biomaterialia. 2007; 3(4):413–438.

19. Olson MF, Sahai E. The actin cytoskeleton in cancer cell motility. Clinical & Experimental Metastasis. 2009; 26(4):273–287.

20. Guck J, Schinkinger S, Lincoln B, Wottawah F, Ebert S, Romeyke M, Lenz D, Erickson HM, Ananthakrishnan R, Mitchell D, Käs J, Ulvick S, Bilby C. Optical deformability as an inherent cell marker for testing malignant transformation and metastatic competence. Biophysical Journal. 2005; 88(5):3689–3698.

21. Moustakas A, Stournaras C. Regulation of actin organisation by TGF-beta in H-ras-transformed fibroblasts. Journal of Cell Science. 1999; 112(Pt 8):1169–1179.

22. Rao JY, Hurst RE, Bales WD, Jones PL, Bass RA, Archer LT, Bell PB, Hemstreet GP. 3rd. Cellular F-actin levels as a marker for cellular transformation: relationship to cell division and differentiation. Cancer Research. 1990; 50(8):2215–2220.

23. Ananthakrishnan R, Menon S. Integration of cell biology, pharmacological modeling and statistical analysis: part I: cell biology and PK/PD in the Oncology paradigm. Critical Reviews in Oncology/Hematology. 2012; 83(2):153–169.

24. Park S, Koch D, Cardenas R, Käs J, Shih CK. Cell motility and local viscoelasticity of fibroblasts. Biophysical Journal. 2005;89(6):4330–4342.

25. Parker AL, Kavallaris M, McCarroll JA. Microtubules and their role in cellular stress in cancer. Frontiers in Oncology. 2014; 4:153.

26. Hernandez P, Tirnauer JS. Tumor suppressor interactions with microtubules: keeping cell polarity and cell division on track. Disease Models & Mechanisms. 2010; 3(5–6):304–315.

27. Kidd ME, Shumaker DK, Ridge KM. The role of vimentin intermediate filaments in the progression of lung cancer. American Journal of Respiratory Cell and Molecular Biology. 2014; 50(1):1–6.

28. Satelli A, Li S. Vimentin in cancer and its potential as a molecular target for cancer therapy. Cellular and Molecular Life Sciences. 2011; 68(18):3033–3046.

29. Hendrix MJ, Seftor EA, Chu YW, Trevor KT, Seftor RE. Role of intermediate filaments in migration, invasion and metastasis. Cancer Metastasis Reviews. 1996; 15(4):507–525. Review.

30. Li J, Wang R, Yang L, Wu Q, Wang Q, Nie Z, Yu Y, Ma J, Pan Q. Knockdown of Nestin inhibits proliferation and migration of colorectal cancer cells. International Journal of Clinical and Experimental Pathology. 2015; 8(6):6377–6386.

31. Aguda BD. Kick-starting the cell cycle: from growth-factor stimulation to initiation of DNA replication. Chaos. 2001; 11(1):269–76.

32. http://www.biology.arizona.edu/CELL_Bio/tutorials/cell_cycle/cells3.html

33. DeVita VT Jr, Chu E. A history of cancer chemotherapy. Cancer Research. 2008; 68(21):8643–8653.

34. https://en.wikipedia.org/wiki/Chemotherapy#Types

35. https://en.wikipedia.org/wiki/Anthracycline

36. Janssen A, Medema RH. Mitosis as an anti-cancer target. Oncogene. 2011; 30(25):2799–2809.

37. Kurzrock R, Markman M. Targeted cancer therapy. Humana Press; 2008.

38. Hudis CA. Trastuzumab – mechanism of action and use in clinical practice. New England Journal of Medicine. 2007; 357(1):39–51.

39. Druker BJ. Imatinib alone and in combination for chronic myeloid leukemia. Seminars in Hematology. 2003; 40(1):50–58. Review.

40. Tonra JR, Hicklin DJ. Targeting the vascular endothelial growth factor pathway in the treatment of human malignancy. Immunological Investigations. 2007; 36(1):3–23. Review.

41. Peters C, Brown S. Antibody-drug conjugates as novel anti-cancer chemotherapeutics. Bioscience Reports. 2015; 35(4). pii: e00225.

42. Martínez MT, Pérez-Fidalgo JA, Martín-Martorell P, Cejalvo JM, Pons V, Bermejo B, Martín M, Albanell J, Lluch A. Treatment of HER2 positive advanced breast cancer with T-DM1: A review of the literature. Critical Reviews in Oncology/Hematology. 2016; 97:96–106.

43. Jordan VC, Furr BJA. Hormone therapy in breast and prostate cancer. Humana Press; 2009.

44. Labrie F. Hormonal therapy of prostate cancer. Progress in Brain Research. 2010; 182:321–341.

45. Wong AC, Ma B. An update on the pharmacodynamics, pharmacokinetics, safety and clinical efficacy of nivolumab in the treatment of solid cancers. Expert Opinion on Drug Metabolism & Toxicology. 2016 Oct;12(10):1255–1261.

46. Bustamante Alvarez JG, González-Cao M, Karachaliou N, Santarpia M, Viteri S, Teixidó C, Rosell R. Advances in immunotherapy for treatment of lung cancer. Cancer Biology & Medicine. 2015; 12(3):209–222.

47. Hughes T, Klairmont M, Sharfman WH, Kaufman HL. Interleukin-2, Ipilimumab, and Anti-PD-1: Clinical Management and the Evolving Role of Immunotherapy for the Treatment of Patients With Metastatic Melanoma. Cancer Biology & Therapy. 2015 Sep 29:0.

48. Sgambato A, Casaluce F, Sacco PC, Palazzolo G, Maione P, Rossi A, Ciardiello F, Gridelli C. Anti PD-1 and PDL-1 immunotherapy in the treatment of advanced non-small cell lung cancer (NSCLC): a review on toxicity profile and its management. Current Drug Safety. 2015 Sep 27.

49. https://www.cancer.org/treatment/treatments-and-side-effects/treatment-types/immunotherapy/monoclonal-antibodies.html

50. Ananthakrishnan R. Gona P. Pharmacological modeling and biostatistical analysis of a new drug. Open Access Journal of Clinical Trials. 2010; 2:59–82.

51. Ananthakrishnan R, Menon S. Design of oncology clinical trials: a review. Critical Reviews in Oncology/Hematology. 2013; 88(1):144–153.

52. Rajman I. PK/PD modelling and simulations: utility in drug development Drug Discovery Today. 2008; 13(7–8):341–346.

53. Gandhi L, Bahleda R, Tolaney SM, Kwak EL, Cleary JM, Pandya SS, Hollebecque A, Abbas R, Ananthakrishnan R, Berkenblit A, Krygowski M, Liang Y, Turnbull KW, Shapiro GI, Soria JC. Phase I study of neratinib in combination with temsirolimus

in patients with human epidermal growth factor receptor 2-dependent and other solid tumors. Journal of Clinical Oncology. 2014; 32(2):68–75.

54. Awada A, Dirix L, Manso Sanchez L, Xu B, Luu T, Diéras V, Hershman DL, Agrapart V, Ananthakrishnan R, Staroslawska E. Safety and efficacy of neratinib (HKI-272) plus vinorelbine in the treatment of patients with ErbB2-positive metastatic breast cancer pretreated with anti-HER2 therapy. Annals of Oncology. 2013; 24(1):109–116.

55. Gomez DL, Armando RG, Cerrudo CS, Ghiringhelli PD, Gomez DE. Telomerase as a Cancer Target. Development of New Molecules. Current Topics in Medicinal Chemistry. 2016; 16(22):2432–2440. Review.

56. Jafri MA, Ansari SA, Alqahtani MH, Shay JW. Roles of telomeres and telomerase in cancer, and advances in telomerase-targeted therapies. Genome Medicine. 2016; 8(1):69.

57. Yoshida GJ, Saya H. Therapeutic strategies targeting cancer stem cells. Cancer Science. 2016; 107(1):5–11.

58. Dragu DL, Necula LG, Bleotu C, Diaconu CC, Chivu-Economescu M. Therapies targeting cancer stem cells: Current trends and future challenges. World Journal of Stem Cells. 2015; 7(9):1185–1201.

59. Mitra AK, Agrahari V, Mandal A, Cholkar K, Natarajan C, Shah S, Joseph M, Trinh HM, Vaishya R, Yang X, Hao Y, Khurana V, Pal D. Novel delivery approaches for cancer therapeutics. Journal of Controlled Release. 2015; 219:248–268.

60. Postel-Vinay,S. Redefining dose-limiting toxicity, Clinical Advances in Hematology & Oncology. 2015; 13(2):87–89.

61. Gelmon,K.A., Eisenhauer, E.A., Harris, A.L., Ratain, M.J., Workman, P., Anticancer agents targeting signaling molecules and cancer cell environment: challenges for drug development? Journal of the National Cancer Institute. 1999; 91(15):281–1287.

62. Sverdlov, O., Wong, W.K., Ryeznik, Y. Adaptive clinical trial designs for phase I cancer studies, Statistics Surveys. 2014; 8:2–44.

63. Braun, T.M., The current design of oncology phase I clinical trials: progressing from algorithms to statistical models, Chinese Clinical Oncology. 2014; 3(1):2.

64. Zang, Y., Lee, J.J. Adaptive clinical trial designs in oncology, Chinese Clinical Oncology. 2014; 3(4):49.

65. Wong, K.M., Capasso, A., Eckhardt, S.G., The changing landscape of phase I trials in oncology, Nature Reviews. Clinical Oncology. 2016; 13(2):106–117.

66. O'Quigley, J., Pepe, M., Fisher, M.L., Continual reassessment method: a practical design for phase 1 clinical trials in cancer, Biometrics. 1990; 46:33–48.

67. Thall, P.F., Cook, J.D., Estey, E. Adaptive dose selection using efficacy-toxicity trade-offs: illustrations and practical considerations, Journal of Biopharmaceutical Statistics. 2006; 16(5):623–638.

68. Hansen, A.R., Graham, D.M., Pond, G.R., Siu,L.L. Phase 1 trial design: is 3 + 3 the best? Cancer Control. 2014; 21(3):200–208.

69. Iasonos, A., Gönen, M., Bosl, G.J., Scientific Review of Phase I Protocols With Novel Dose-Escalation Designs: How Much Information Is Needed? Journal of Clinical Oncology. 2015; 33(19):2221–2225.

70. Jaki, T., Clive, S., Weir, C.J., Principles of dose finding studies in cancer: a comparison of trial designs, Cancer Chemotherapy and Pharmacology. 2013; 71(5):1107–1114.

71. Eisenhauer, E.A., O'Dwyer, P.J., Christian, M., Humphrey, J.S. Phase I clinical trial design in cancer drug development, Journal of Clinical Oncology. 2000; 18(3): 684–692.

72. Ji, Y., Liu, P., Li, Y., Bekele, B.N. A modified toxicity probability interval method for dose-finding trials, Clinical Trials. 2010; 7(6):653–663.

73. Blanchard, M.S., Longmate, J.A. Toxicity equivalence range design (TEQR): a practical Phase I design, Contemporary Clinical Trials. 2011; 32(1):114–121.

74. Liu, S., Yuan, Y. Bayesian optimal interval designs for phase I clinical trials, Journal of the Royal Statistical Society: Series C. 2015; 64:507–523.

75. Babb, J.S., Rogatko, A. Patient specific dosing in a cancer phase I clinical trial, Statistics in Medicine. 2001; 20(14):2079–2090.

76. Tighiouart, M., Rogatko, A. Dose Finding with Escalation with Overdose Control (EWOC) in Cancer Clinical Trials, Statistical Science. 2010; 25(2): 217–226.

77. Yang, S., Wang, S.J., Ji, Y. An integrated dose-finding tool for phase I trials in oncology, Contemporary Clinical Trials. 2015: 45(Pt B):426–434.

78. Storer, B.E. An evaluation of phase I clinical trial designs in the continuous dose–response setting, Statistics in Medicine. 2001; 20:2399–2408.

79. Simon, R., Freidlin, B., Rubinstein, L., Arbuck, S.G., Collins, J., Christian, M.C. Accelerated titration designs for phase I clinical trials in oncology, Journal of the National Cancer Institute. 1997; 89(15):1138–1147.

80. Ivanova, A. Escalation, group and A + B designs for dose-finding trials, Statistics in Medicine. 2006; 25(21):3668–3678.

81. Le Tourneau, C., Lee, J.J., Siu, L.L. Dose escalation methods in phase I cancer clinical trials, Journal of the National Cancer Institute. 2009; 101(10):708–720.

82. Liu, S., Cai, C., Ning, J. Up-and-down designs for phase I clinical trials, Contemporary Clinical Trials. 2013; 36:218–227.

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

84. Garrett-Mayer, E. The continual reassessment method for dose-finding studies: a tutorial, Clinical Trials. 2006; 3(1):57–71.

85. Lee, S.M., Cheung, Y.K. Model calibration in the continual reassessment method, Clinical Trials. 2009; 6:227–238.

86. Paoletti, X., Baron, B., Schöffski, P., Fumoleau, P., Lacombe, D., Marreaud, S., Sylvester, R. Using the continual reassessment method: lessons learned from an EORTC phase I dose finding study, European Journal of Cancer. 2006; 42(10): 1362–1368.

87. Ivy, S.P., Siu, L.L., Garrett-Mayer, E., Rubinstein, L. Approaches to phase 1 clinical trial design focused on safety, efficiency, and selected patient populations: a report from the clinical trial design task force of the national cancer institute investigational drug steering committee, Clinical Cancer Research. 2010; 16(6): 1726–1736.

88. Ivanova, A., Montazer-Haghighi, A., Mohanty, S.G., Durham, S.D. Improved up-and-down designs for phase I trials, Statistics in Medicine. 2003; 22(1):69–82.

89. Ivanova, A., Dose-Finding in Oncology – Nonparametric Methods (Chapter 4), in: N. Ting, (Editor), Dose Finding in Drug Development. Springer, New York, 2006, pp. 49–58

90. Le Tourneau, L., Gan, H.K., Razak, A.R., Paoletti, X., Efficiency of new dose escalation designs in dose-finding phase I trials of molecularly targeted agents, PLoS One. 2012; 7(12):e51039.

91. Huang B, Bycott P, Talukder E. Novel dose-finding designs and considerations on practical implementations in oncology clinical trials. Journal of Biopharmaceutical Statistics. 2017; 27(1):44–55.

92. Sato H, Hirakawa A, Hamada C. An adaptive dose-finding method using a change-point model for molecularly targeted agents in phase I trials. Statistics in Medicine. 2016; 35(23):4093–4109.

93. Wages NA, Tait C. Seamless Phase I/II Adaptive Design for Oncology Trials of Molecularly Targeted Agents. Journal of Biopharmaceutical Statistics. 2015; 25(5):903–920.

94. Markman M. Serious ethical dilemma of single-agent pegylated liposomal Doxorubicin employed as a control arm in ovarian cancer chemotherapy trials. Journal of Clinical Oncology. 2010 Jul 1; 28(19):e319–320.

95. Ferrandina G. Reply to M. Markman, Journal of Clinical Oncology. 2010 Jul 1; 28(19):e321–22.

96. Schilsky RL, Minasian L, Auclair D, Rahman A, Pazdur R., Optimizing Dosing of Oncology Drugs. Conference on Clinical Cancer Research. 2013.

97. Pan H, Huang P, Wang Z, Wang L, Li C, Xia J. A novel Bayesian seamless phase I/II design. PLoS One. 2013 Sep 4; 8(9):e73060.

98. Pan H, Xie F, Liu P, Xia J, Ji Y. Clin Trials. A phase I/II seamless dose escalation/expansion with adaptive randomization scheme (SEARS). Clinical Trials. 2014; 11(1):49–59.

99. Hoering A, LeBlanc M, Crowley J. Seamless phase I–II trial design for assessing toxicity and efficacy for targeted agents. Clinical Cancer Research. 2011 Feb 15; 17(4):640–646.

100. Hoering A, Mitchell A, LeBlanc M, Crowley J. Early phase trial design for assessing several dose levels for toxicity and efficacy for targeted agents. Clinical Trials. 2013; 10(3):422–429.

101. Zang Y, Lee JJ, Yuan Y. Adaptive designs for identifying optimal biological dose for molecularly targeted agents. Clinical Trials. 2014 Jun; 11(3):319–327.

102. Thall PF, Cook JD. Dose-finding based on efficacy-toxicity trade-offs. Biometrics. 2004; 60(3):684–693.

103. Braun TM. The bivariate continual reassessment method. extending the CRM to phase I trials of two competing outcomes. Controlled Clinical Trials. 2002 Jun; 23(3):240–256.

104. Yin G, Li Y, Ji Y. Bayesian dose-finding in phase I/II clinical trials using toxicity and efficacy odds ratios. Biometrics. 2006 Sep; 62(3):777–784.

105. Zhang W, Sargent DJ, Mandrekar S. An adaptive dose-finding design incorporating both toxicity and efficacy. Statistics in Medicine. 2006 Jul 30; 25(14):2365–2383.

106. Dragalin V, Fedorov VV, Wu Y. Two-stage design for dose-finding that accounts for both efficacy and safety. Statistics in Medicine. 2008 Nov 10; 27(25):5156–5176.

107. Yuan Y, Nguyen HQ, Thall PF. Bayesian Designs for Phase I–II Clinical Trials. Chapman and Hall/CRC. 2016.

108. Ananthakrishnan R, Green S, Chang M, Doros G, Massaro J, LaValley M, Systematic Comparison of the Statistical Operating Characteristics of Various Phase 1 Oncology Designs. Contemporary Clinical Trials Communications. 2017; 5:34–48.

109. Ivanova A, Liu K, Snyder E, Snavely D. An adaptive design for identifying the dose with the best efficacy/tolerability profile with application to a crossover dose-finding study. Statistics in Medicine. 2009 Oct 30; 28(24):2941–2951.

110. Demirtas H, Hedeker D. A practical way for computing approximate lower and upper correlation bounds. The American Statistician. 2011; 65(2):104–109.

111. Mullard A. Reining in the supersized Phase I cancer trial. Nature Reviews. Drug Discovery. 2016 Jun 1; 15(6):371–373.

112. Iasonos A, Zohar S, O'Quigley J. Incorporating lower grade toxicity information into dose finding designs. Clinical Trials. 2011 Aug; 8(4):370–379.

113. Gezmu M, Flournoy N. Group up-and-down designs for dose finding. Journal of Statistical Planning and Inference. 2006; 136:1749–1764.

114. Li DH, Whitmore JB, Guo W, Ji Y. Toxicity and Efficacy Probability Interval Design for Phase I Adoptive Cell Therapy Dose-Finding Clinical Trials. Clinical Cancer Research. 2017 Jan 1; 23(1):13–20.

115. Tibshirani R, Hoefling H, Tibshirani R. Nearly-Isotonic Regression. Technometrics. 2011; 53(1):54–61.

116. Braun TM. Generalizing the TITE-CRM to adapt for early- and late-onset toxicities. Statistics in Medicine. 2006 Jun 30; 25(12):2071–2083.

117. Polley MY. Practical modifications to the time-to-event continual reassessment method for phase I cancer trials with fast patient accrual and late-onset toxicities. Statistics in Medicine. 2011 Jul 30; 30(17):2130–2143.

118. Chen M. Dose Finding Methods Based on Cure Rate Model in Phase I Cancer Clinical Trials. Dissertation Submitted to Rutgers. 2015.

119. Bekele BN, Shen Y. A Bayesian approach to jointly modeling toxicity and biomarker expression in a phase I/II dose-finding trial. Biometrics. 2005 Jun; 61(2):343–354.

VITA