Early phase clinical trials are essential in the drug development process as they provide key information about new interventions which can be used in later-phase testing. Therefore, it is important that the decisions made and conclusions reached in the early phase setting are correct or as accurate as possible. Failure to do so could lead to a waste of resources in pursuing unnecessary further research and could potentially negatively impact patients well-being. This thesis aims to explore methodologies used in early phase dose-finding trials. We look specifically at the implementation of novel methods as well as how they can be extended for use in complex and innovative designs.

In this chapter, we begin with a brief introduction of clinical trials and also provide a summary of the underlying methodology used within this thesis specific to early phase trials. A description of each chapter is then provided.

Clinical trials are often a time-consuming and costly process \cite{fogelFactorsAssociatedClinical2018}. It can take 10-20 years to get a new drug from inception to regulatory approval \cite{lipskyIdeaMarketDrug2001, mohsDrugDiscoveryDevelopment2017}. Any new treatments that come to market must be thoroughly tested and examined, to make sure not only that they are safe but also effective and better than treatments currently in use.

The clinical trial process is split into multiple stages, commonly referred to as phases. Each phase has a different objective and builds upon knowledge and data collected in the previous phases \cite{wrightChapterClinicalTrial2017}. Phase I trials aim to determine the safety of a treatment. This typically takes the form of a dose-finding study that aims to find a safe and tolerable dose that can be taken forward for further testing \cite{iasonosRandomisedPhaseClinical2021}. Phase II trials aim to determine if a treatment works or if there is some signal of efficacy. This is usually done with a single arm design i.e. a sample of patients is given the experimental treatment \cite{esteyNewDesignsPhase2003}. However, some Phase II trials can be RCTs (Randomised Controlled Trials), where the new treatment is compared with a treatment already in use or a placebo \cite{mandrekarRandomizedPhaseII2010}. Finally, Phase III trials aim to establish the efficaciousness of an experimental treatment. Normally, this is done by comparing the new treatment against the standard of care. Results from Phase III clinical trials can then go on to

influence clinical practice \cite{umscheidKeyConceptsClinical2011}.

We usually consider Phase I and single arm Phase II trials to be early phase trials, then randomized Phase II and Phase III trials as late phase trials. However, there is no definitive boundary between early and late phase trials. Phases are determined by their intention and not by their design.

For Phase \RN{1} trials the main aim is to establish a dose, commonly referred to as the maximum tolerated dose (MTD), which can then be taken forward into later phase testing. Normally these may be considered first-in-human trials as Phase \RN{1} is typically the point at which a drug would first be tested in healthy human volunteers. However, in the oncology setting this is generally not the case. Often times, due to the nature of the treatments, such as chemotherapy or radiotherapy, they may be considered too toxic to give to healthy individuals so are rather tested in patients with the specific disease of interest \cite{salzbergFirstinHumanPhaseStudies2012}.

Traditionally these trials adopted algorithm-based approaches. Here predetermined rules were used during the trial to allocate patients to dose-levels and select the MTD. An example of this would be the 3+3 design \cite{storerDesignAnalysisPhase1989}, where patients were recruited in cohorts of three and dependent on the outcomes observed in each cohort a specific decision on the next investigative dose would be made. There were many criticisms of these approaches as they resulted in the sub-optimal treatment of patients, poor operating characteristics and a recommended dose or MTD that has limited interpretation as a dose yielding a specific target toxicity \cite{iasonosComprehensiveComparisonContinual2008, onar-thomasSimulationbasedComparisonTraditional2010}.

This gave rise to the continual reassessment method (CRM) which was first introduced by O'Quigley et al. \cite{oquigleyContinualReassessmentMethod1990} in 1990. This methodology was developed as an approach to meet ethical requirements and use models to reasonably approximate the true probability of toxicity around the dose close to some pre-defined target toxicity. In their paper, O'Quigley et al. \cite{oquigleyContinualReassessmentMethod1990} demonstrate the CRM's superiority over various algorithm based designs via simulations. The main advantage of the CRM is that it is able to make use of all accumulated data whereas designs such as the 3+3 make decisions and recommendations based on data from the most recent cohort of patients. In the case of the 3+3 design, escalation continues until at least two patients in a cohort of three or six experience a DLT. More explicitly, the MTD is the dose level below the dose at which $\geq$33\% of patients experience a DLT \cite{letourneauDoseEscalationMethods2009}. %\cite{letourneauDoseEscalationMethods2009}.

With the CRM debuting over 30 years ago in the literature and multiple papers over the subsequent years confirming its advantages over rule-based designs you would expect model-based approaches for dose-finding trials to become the norm however, this is not the case. A study by Rogatko et al. \cite{rogatkoTranslationInnovativeDesigns2007} published in 2007 looked into the translation of effective statistical designs into phase \RN{1} trials for new anticancer therapies. Between 1991 and 2006 they searched for abstracts and categorised them as either clinical dose-finding trials or statistical methodology for dose-escalation trials. They found 1235 clinical trials and 90 methodological papers. Of those 1235 trials only 20 (1.6\%) used statistical methodology, the remaining papers used various rule-based designs. A later review by Chiuzan et al. \cite{chiuzanDosefindingDesignsTrials2017} looked at the number of phase \RN{1} oncology articles published between 2008 and 2014. Out of the 1712 dose-finding trials 1591 (92.9\%) used rule-based designs.

Based on these reviews we can see that the uptake of more efficient model-based designs such as the CRM has been slow and limited. There are probably a number of factors which cause this, such as lack of resources, access and understanding. The main issue is that implementation of these designs usually require the input of a statistician, more specifically one who is familiar with such approaches. They also need to be able to implement and conduct the trial with software available, for designs such as the CRM there are multiple options available such as the R packages dfcrm \cite{cheungDfcrmDoseFindingContinual2019}, escalation \cite{brockModularApproachDose2020} and trialr \cite{brockTrialrClinicalTrial2023}. However, for more complex and innovative designs, software may not be readily accessible and implementation may be difficult.

Phase \RN{2} trials build upon the work of early Phase \RN{1} trials. Here the focus shifts away from toxicity and looks more towards efficacy of these new treatments at the dose-levels previously determined in phase \RN{1} trials \cite{berryBayesianAdaptiveMethods2010}. The key purpose of phase \RN{2} trials is to see if a new treatment or intervention works and establish if there is an efficacy signal. More specifically they aim to determine if there is a sufficient level of efficacy to warrant further research in for example a Phase \RN{3} setting \cite{juliousIntroductionStatisticsEarly2010}.

An example of an early Phase \RN{2} trial design would be a single arm trial. Eligible patients come into the trial

and all of them will be allocated to the new treatment. Once they have completed their treatment period we would then assess the effectiveness of that treatment using some measure of success. Looking at the outcome of success in each patient, the success rate or proportion of success can then be determined. In the single arm setting, this success rate is then compared to some sort of benchmark, which is determined from either historical data or clinical experience.

One approach to analysing a trial like this is Bayesian and utilises a Beta-Binomial conjugate analysis to estimate a response rate for a binary outcome. This is a fairly straight-forward analysis and can be found in most Bayesian text books, one example of which is by Lee \cite{leeBayesianStatisticsIntroduction2012}. Whilst this may be somewhat simplistic from a mathematical standpoint, clinicians may be less familiar with Bayesian approaches in general compared to frequentist methods. However, there is still value in using these methods to analyse trials as they can often be more efficient especially with smaller sizes \cite{inoueRelationshipBayesianFrequentist2005}, which is often the case in an early phase setting and they allow for greater flexibility. Using a Bayesian approach allows for decisions to be made based on probabilities from a posterior distribution \cite{savilleUtilityBayesianPredictive2014}. Which can be a more intuitive way to understand the treatment effect given observed data.

These types of analyses also better facilitate other complex and innovative designs such as platform/basket/umbrella trials where multiple biomarkers or treatments may be under investigation. The Bayesian approach will allow for information borrowing across any different arms in the trial, which is particularly useful when working with restricted resources \cite{carlinBayesianComplexInnovative2022}.

Generally speaking, early phase trials work with less resources (i.e less patients, time, money), so it would be advantageous to use designs which are more efficient with the data collected, the majority of which would require a statistician familiar with these methods to implement. Clinicians may also push for designs that are easier to follow so either a algorithm-based approach in a Phase \RN{1} or a frequentist approach for a Phase \RN{2} trial. Clinicians may also not be familiar with these complex designs and how they work.

As developments occur in the medical field, e.g. introducing new interventions, our trials may subsequently become more complex. This complexity arises because the assumptions made by current methodologies may no longer hold true. Additionally, the questions that trials aim to answer can also become more intricate due to various factors, including the disease setting under investigation and the type of treatment involved. Consequently, it is essential to address these challenges as they arise to ensure the optimal utilisation of our data and to make accurate decisions.

Whilst the development of new methodologies is necessary to facilitate better research, practical considerations also need to be made. The ease at which a new methodology can be implemented is important. Whilst a new method may be fantastic and solve a lot of issues, if researchers cannot easily implement it either through software or replicable code then the method will ultimately have little impact in the real world. Similarly, communication of new methods is also important. Not only for other statisticians but the plethora of multidisciplinary teams involved through the trial process. A summary of each chapter is provided in the next section, where we address different aspects of this further.

In Chapter \ref{Adept}, we detail our experiences implementing a novel methodology into the design of an early phase dose-finding trial in head and neck cancer. This trial, ADePT-DDR run by the University of Birmingham (UoB) Cancer Research Clinical UK Trials Unit (CRCTU), uses the partial ordering time-to-event continual reassessment method (PO-TITE-CRM) design \cite{wagesUsingTimetoeventContinual2013}. The PO-TITE-CRM design was introduced in 2013 as an extension to the TITE-CRM design, itself an extension of the original continual reassessment method (CRM), a model-based approach to dose-finding trials. Despite the publication of this novel dose-escalation design its implementation appears to be rare. The CRM operates under the monotonicity assumption which assumes that as the dosage of drug increases so does the probability of toxicity. The PO-TITE-CRM was developed to address scenarios where the ability to fully order the doses based on increasing toxicity is missing. Multiple iterations of simulations were utilised to determine the optimal parameterisation of the design. Simulation results from the optimal parameterisation show the operating characteristics of the design perform well across a variety of scenarios. Further work was then done to compare the design we chose against potential alternatives. We present an overview of the design methodology and its application in this trial scenario.

The focus of Chapter \ref{WT} is on an extension to a seamless Phase \RN{1}/\RN{2} design. This design by Wages and Tait \cite{wagesSeamlessPhaseII2015} uses adaptive randomisation to conduct its dose-finding. We aimed to extend this design to include a control comparator arm by leveraging the design's adaptive randomisation mechanism. Typically these types of seamless design utilise both toxicity and efficacy outcomes to select a dose for later phase testing. Our motivation was to take this one step further and develop a seamless Phase \RN{1}/\RN{2} design that would allow for a direct comparison of the selected dose to a control arm. We present the modification we make to the design and detail how altering specific parameters impacts the allocation of patients to the control arm within the design. Simulations were then conducted to investigate the operating characteristics under certain specifications. Further simulation work is then used to compare our modified design with the original Wages and Tait design to ascertain if performance is impacted by our modification. Additionally, this control arm allows us to conduct power calculations comparing efficacy rates between patients allocated to the optimal dose and those in the control arm. These calculations give an insight into how our design would perform as a standard Phase \RN{2} trial.

Decision-making is an important component of dose-finding trials. For non-statisticians, understanding the reason why certain decisions are being made from a model-based approach such as the CRM could be confusing or unintuitive compared to a 3+3 for example. In order to bridge that gap, Yap et al. developed a novel visualisation tool called dose transition pathways (DTPs) \cite{yapDoseTransitionPathways2017}. Our work in Chapter \ref{TITE-DTP}, explores the use of DTPs in a time-to-event (TITE) setting. TITE methodology, in a dose-finding context, allows decisions to be made at earlier time points based on partially observed data from patients currently in a trial. This leads to the number of outcomes being more complicated for a single cohort of patients in a dose-finding trial. This in turn further complicates DTPs which aim to effectively summarise the different dose-decisions that can be made based on all possible outcomes. Our work aims to reap the benefits of DTPs in a TITE setting. We detail what those benefits are and the challenges in presenting DTPs for TITE methodology. Using informative examples we demonstrate what these challenges are and why they occur as well as potential solutions for how DTPs can still be incorporated for trials utilising this methodology.

In Chapter \ref{etp}, we introduce efficacy transition pathways (ETPs) which are a new visualisation tool inspired by DTPs for use in Phase \RN{2} trials. The idea of visualising outcomes and decisions for dose-finding trials is a useful one, not only during the conduct of a trial but in its initial design stages as well. Through the development of three different trials at CRCTU it became apparent that a similar tool would be beneficial for Phase \RN{2} trials. These trials, whilst not conducting dose-finding, aimed to assess efficacy and included multiple interim analyses at which different decisions could be made. We describe how ETPs are constructed and work, by providing examples from trials which inspired the idea behind ETPs. One of the success of DTPs was its ease of implementation through available software like the R package escalation by Brock \cite{brockModularApproachDose2020}. In order to facilitate the same ease of use for ETPs we developed a R function along with a web based application to automatically generate these plots. In the chapter we detail how the application works to construct ETPs and additional features that we added so it could be used as an educational tool.

To end, Chapter \ref{Conclusion}, provides an overarching summary of the topics and ideas discussed in this thesis.