Indian Statistical Institute

Doctoral Dissertation

Determination of Sample Size in Vaccine Trials

A thesis submitted to the Indian Statistical Institute in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Statistics

Author: **MEGHNA BOSE**

Supervisor: **PROF. ATANU BISWAS**

Applied Statistics Unit (ASU)

Statistical Sciences Division

Abstract

"Brevity is the soul of wit."

— William Shakespeare, Hamlet.

Sample size determination is an important as well as delicate problem in clinical trials. Often, prior information available before the beginning of the trial is used to estimate the sample size. Hence, it is of importance to quantify the prior information available in terms of an Effective Sample Size (ESS). The problem of finding ESS is mainly done in Phase-II trials, where the responses are usually bivariate—one variable corresponds to toxicity and the other corresponds to efficacy. The two variables can be binary, continuous, or one of each type. Since there is already an existing literature on finding ESS when the two variables are of the same type (both binary or continuous), we focussed on finding ESS when one of the variables is binary and the other is continuous. The problem is statistically more challenging in case of mixed responses, and we propose a novel method for finding ESS in these situations. Without loss of generality, we explored the case of binary efficacy and continuous safety for different prior distributions under different setups. Theoretical expressions have been obtained in various situations. The methods have been evaluated and compared by simulation studies. The proposed method is then illustrated by using some real-life data on a Phase II vaccine trial for COVID-19.

Once the problem of finding ESS was solved, we shifted our attention to the actual problem of sample size determination, which basically gives an idea about the number of individuals to be recruited in the Phase-III trials. The emergence of the COVID-19 pandemic sparked public interest in vaccines. It was then we realized that the sample sizes in vaccine trials deserve equal attention as that of the usual drug trials.

The Phase-III trial of a vaccine recruits a group of individuals, of whom some are given the vaccine and the rest are allocated to a placebo (or control) group. The individuals in both the groups are then followed till a fixed period of time and checked for the occurrence of the disease. The effectiveness of a vaccine in preventing a particular disease is measured by means of protective vaccine efficacy, defined as $VE = 1 - \frac{ARV}{ARU}$, where ARV and ARU are respectively the disease attack rates in the vaccinated and unvaccinated population. For each of the cohort and case-control designs, the sample sizes are

calculated by fixing the maximum width of the $100(1-\alpha)\%$ confidence interval (CI) of VE and finding the minimum sample size required to achieve that width. The majority of the literature focuses on obtaining sample sizes when the equal number of individuals are allocated across the vaccine and placebo groups. We present methods for calculating the required sample size with a specified degree of precision when there is an unequal allocation of individuals across the two groups. The fraction of individuals allocated to the placebo group (ρ) is chosen in such a way that some relevant criterion of interest is optimized. Two such criteria considered are the minimization of the total sample size and the expected number of people developing the disease, and the resultant optimal allocation proportions, respectively, give rise to Neyman and RSIHR allocations. It was observed that the sample sizes and the optimal allocation proportions come out to be a function of the parameters VE and ARU, which are unknown and to be estimated from the trial. Therefore, we present strategies for implementing the optimal allocations in real-life trials through sequential methods such as sequential maximum likelihood estimation (SMLE), doubly biased coin design (DBCD), and efficient randomized-adaptive designs (ERADE) procedures. However, certain drawbacks in implementation has been pointed out, and the aforementioned procedures help us achieve Neyman and RSIHR allocation proportions asymptotically but do not provide an estimate of the sample size at the beginning of the trial. Hence, one has to rely on the guesses about the unknown parameters (usually available from Phase II or some previously conducted trial) to get an insight about the sample size. It is quite evident that wrong or uncertain guesses about the values will give misleading results. The Bayesian paradigm comes to the rescue in such a case, which helps us deal with the uncertainties associated with the assumed values of the unknown parameters. This motivated us towards our third work, which provides a Bayesian alternative to the sample size determination problem in vaccine efficacy (both cohort and case-control) studies. Instead of the CI, the Bayesian approach considers HPD/credible interval-based criteria, viz. Average Coverage Criteria (ACC) and Average Length Criteria (ALC), to determine the sample sizes. Optimal allocations analogous to Neyman and RSIHR has been developed in the Bayesian case as well. The Bayesian sample sizes came out to be a function of the hyper-parameters of the priors taken on ARV and ARU, and by choosing the values of the hyper-parameters suitably, we can deal with the uncertainties associated with their assumed values. Both the frequentist and Bayesian sample sizes were compared among themselves and with those taken in a real-life trial.

It has already been mentioned that the participants in both vaccine and placebo groups are monitored for a fixed period of time and number of cases of the disease developed at the end of the follow-up period is noted. The proportion of people developing the disease at the end of the follow-up in both groups serves as estimates of ARV and ARU. One assumes that the number of cases of the disease developed in the vaccinated and unvaccinated groups is binomially distributed with success probability parameters ARV

and ARU, respectively. However, it is to be noted that the estimates of ARV, ARU, and subsequently that of VE, vary with the change in time to follow-up. Moreover, the datasets obtained from vaccine trials are of time-to-event (time to development of the disease) nature, and survival models give a better fit than the conventionally used binomial model. In the survival set-up, the vaccine efficacy is redefined as $VE_H(t) = 1 - \frac{\lambda_B(t)}{\lambda_A(t)}$, where $\lambda_A(t)$ and $\lambda_B(t)$ are respectively the hazard rates of the unvaccinated and vaccinated population at time t. We then approach the sample size determination problem under the survival framework after presuming appropriate distributions on the survival and censoring times. A semi-parametric approach using the Cox Proportional Hazards model, which does not require any distributional assumptions on the survival times, has also been considered. Sample sizes and optimal allocation proportions have been obtained, and their superiority over the binomial sample sizes has been pointed out.

Acknowledgements

Date: January 19, 2026 Name: Meghna Bose

The popular saying goes, *"It's the journey that matters, not the destination."* Rightfully so, when I look back at my PhD journey of 6 years, the memories of people who shaped me into my present self overshadow the reminiscence of the hurdles I faced while building this thesis. These are the people who I look up to, the people who were my support in the face of difficulties, and, most importantly, the people who made me laugh. I would like to acknowledge all these people at the beginning of this thesis and give them their long-due credit.

First of all, I am profoundly thankful to my supervisor, Prof. Atanu Biswas, without whom this thesis would not have been what it is today. He has been a constant source of motivation and helped me grow into the independent researcher I am today (if at all I am!). My sincerest gratitude to my co-author Prof. Jean François Angers, who has been a diligent guide and mentor in all of our joint works. I am extremely fortunate to have the opportunity to learn and grow under the shadow of Prof. Arijit Chakrabarti, Prof. Tapas Samanta, Prof. Sourabh Bhattacharya, Prof. Kiranmoy Das, Prof. Subir Bhandari, and Prof. Abhik Ghosh during my PhD coursework. The knowledge I gained from their courses garnered my interest in the subject and helped me immensely during my doctoral journey. I am truly indebted to the Director of ISI, present and former Deans , RFAC chairs, RFAC conveners, ASU and Dean's Office staff members for smoothly handling the administrative matters.

They say, *"In every conceivable manner, the family is link to our past, bridge to our future."* I am grateful for having a family that is deeply rooted in its past and continuously strives to build a better future for its upcoming generations. I am blessed to have cousins who always have my back and believe in me even if the whole world doesn't. A special thanks to my Mama, who was one of the first people to tell me that Mathematics can actually be pursued as a career option. Another honorary mention must be given to my nephew, Debayan, who, despite being the youngest member, ties the whole family to one place. *I love you, Bablu!!.* Last but not least, I would like to thank my parents for

molding me into the person I am today. My mother is my first teacher and the person who not only introduced me to the world of Mathematics but also got me interested in it. My father is a person who played a huge role in nurturing my ideologies and shaping my mind in the right direction. His in-depth knowledge about every sphere of life amazes me to date. Thank you, Ma and Baba, for everything!! *I dedicate this thesis to both of you.*

It would be unfair if I only talked about the family that was handed to me by birth and not the one I chose for myself—*my friends!!*. I would like to request everyone to hold your horses, as the list is going to be long.

One of the most beautiful memories at ISI is the innumerable gossips, laughters, and witty banters shared at Lakkhan Da's, Sanjib Da's, and subsequently at *Aamtala*. Thanks to my labmates Anik, Sayan, Chirayata, Sayantan, Biswadeep, Javed Da, Soutik, Sampurna, Monitirtha, Amarnath, Arijit, Anirban, Subhankar, Mriganka, Arnab, Sourav, Basak Da, Sujay Da, Rahul, Upama, Annesha, and Chiara for making the tea sessions so memorable. Shout out to Lakkan Da and Sanjib Da as well for providing us with the appropriate platform. Those 'insightful' exchanges will be etched in my heart forever. Life after lab would have been quite mundane without the company of my hostelmates Suman, Gourab, Bibhuti Da, Subha, Anup, Sankar Da, Faizan, Anish, Dipranjan, Priti, Suparna, Jhelum, Soham, and Kirandeep. My friends outside ISI have been a constant source of inspiration as well. My immense thanks to my childhood friends, Sudeshna and Avijit, for always being there for me and providing me with a Premium QuillBot subscription for *free*. From the bottom of my heart, I extend my thanks to my college friends and trek buddies, Snigdha and Shreeja. To many more treks together!! Sonakhya and Chiranjib — Thanks for being in touch even if we don't meet every day. I hope our feminist discussions extend beyond metro rides and span across horizons, and we accompany each other to many more future concerts (especially Teddy's concerts). My heart goes out to my friends whose weddings I attended in the past few years and those I'll be attending in the future. Your weddings were an excellent place for uniting with old friends and felt like a breath of fresh air.

I am truly fortunate to have met quite a few excellent teachers throughout my school, college and university life. It is with great appreciation that I thank each and every one of them as they were the ones who motivated me ti pursue PhD. Also, there have been some non-living entities that have aided me in my journey in more ways than possible. They include ChatGPT, Gemini, DeepSeek, Google, QuillBot, Grammarly, tablesgenerator.com, R Studio, and Overleaf.

Finally, I am honored to thank the Indian Statistical Institute (ISI) for providing me with an exceptional research environment and the most marvelous set of colleagues. Today, with a heavy heart, I express my gratitude to ISI for giving me the *best six years of my life!!*

Some Important Notations and Abbreviations

ESS	Effective Sample Size
ARV	Disease attack rate in the vaccinated population
ARU	Disease attack rate in the unvaccinated population
VE	Vaccine Efficacy
RR	Relative Risk Ratio
OR	Odds Ratio
CI	Confidence Interval
HPD	Highest Posterior Density
RAR	Response Adaptive Randomization
MLE	Maximum Likelihood Estimate
z_α	Upper α point of standard Normal, $N(0,1)$ distribution
$\Phi(.)$	Cumulative Distribution function (CDF) of $N(0,1)$
$\delta(L)$	Degenerate distribution at L
$\mathbb{R}$	Set of Real numbers
$\mathbb{N}$	Set of Natural numbers

Some Probability Densities

The notations used in this thesis for the standard probability distributions along with their probability density (or mass) functions (pdf/pmf), $f(x|.)$, are given below for quick reference. For example, if it is stated that $X \sim N(\mu, \sigma^2)$, it means that X had pdf

$$f(x \mid \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}},\ x \in \mathbb{R},\ \mu \in \mathbb{R},\ 0 < \sigma < \infty.$$

$N(\mu, \sigma^2)$	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}},\ x \in \mathbb{R},\ \mu \in \mathbb{R},\ 0 < \sigma < \infty.$
$Beta(a, b)$	$x^{a-1}(1-x)^{b-1},\ 0 < x < 1,\ a > 0,\ b > 0$
$N_p(\boldsymbol{\mu}, \Sigma)$	$\frac{1}{(2\pi)^{\frac{p}{2}} \lvert\Sigma\rvert^{\frac{1}{2}}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})},\ \boldsymbol{x} \in \mathbb{R}^p,\ \boldsymbol{\mu} \in \mathbb{R},\ \Sigma$ is positive-definite.
$T_p(d, \boldsymbol{\mu}, \Sigma)$	$\frac{\Gamma(\frac{d+p}{2})}{\Gamma(\frac{d}{2}).(d\pi)^{\frac{p}{2}}.\lvert\Sigma\rvert^{\frac{1}{2}}} \cdot \frac{1}{\left\{1 + \frac{(\boldsymbol{x}-\boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}{d}\right\}^{\frac{d+p}{2}}},\ \boldsymbol{x} \in \mathbb{R}^p,$ $\boldsymbol{\mu} \in \mathbb{R},\ \Sigma$ is positive-definite.
$IG(c, d)$	$\frac{d^c}{\Gamma(c)} e^{-\frac{d}{x}} x^{-c-1},\ x > 0,\ c > 0,\ d > 0.$
$Bin(n, p)$	$\binom{n}{x} p^x (1-p)^{n-x},\ x = 0, 1, \ldots, n,\ n \in \mathbb{N},\ 0 < p < 1.$
$Poisson(\lambda)$	$e^{-\lambda} \frac{\lambda^x}{x!},\ x = 0, 1, 2, \ldots, \infty, \lambda > 0.$
$U[a, b]$	$\frac{1}{b-a},\ a < x < b, -\infty < a < b < \infty.$
$Exp(\lambda)$	$\lambda e^{-\lambda x},\ x > 0,\ \lambda > 0.$
$Weibull(\lambda, p)$	$\lambda p x^{p-1} e^{-\lambda x^p},\ x > 0,\ \lambda > 0,\ p > 0.$

List of Publications/Preprints

Certain parts of **Chapter 1** has been written with the help of the *book chapter* —
Bose, M., and Biswas, A. (2025). Overview on Vaccine Efficacy, An. In: Lovric, M. (eds) *International Encyclopedia of Statistical Science*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-69359-9_21.

Chapter 2 is based on the *article* —
Bose, M., Angers, J.F., and Biswas, A. (2022). Prior effective sample size in phase II clinical trials with mixed binary and continuous responses. *Statistica Neerlandica*, 77(2), 233–248. https://doi.org/10.1111/stan.12283.

Chapter 3 is based on the *article* —
Bose, M., and Biswas, A. (2023). Sample sizes required to estimate the protective efficacy of a vaccine when there is an unequal allocation of individuals across the vaccine and placebo groups. *Statistical Methods in Medical Research*, 32(10), 1859-1879. https://doi.org/10.1177/09622802231176807.

Chapter 4 is based on the *preprint* —
Bose, M., Angers, J.F., and Biswas, A. (2024). A Bayesian approach to determining sample sizes in vaccine trials. [*Unpublished manuscript*].

Chapter 5 is based on the *preprint* —
Bose, M., and Biswas, A. (2025). Determining sample sizes after identifying the time-to-event nature of vaccine trial datasets. [*Unpublished manuscript*].

Contents

List of Figures

List of Tables

Chapter 1

Introduction

"To consult the statistician after an experiment is finished is often merely to ask him to conduct a post mortem examination. He can perhaps say what the experiment died of."

— Ronald A. Fisher

Before a drug or vaccine is licensed for use, it has to prove its worth by going through the three phases of a randomized controlled trial. The first phase is a rudimentary stage, where the primary concern is the evaluation of the safety of the drug or vaccine and not its efficacy. Dose-response studies are performed here to determine the optimal dose level of the particular drug or vaccine that causes no serious side effects and can safely be applied to individuals. The subsequent Phase II trial is another small-scale investigation in which the effectiveness of the drug is also considered, even if safety is still an important concern. Hence, the responses at this stage are usually bivariate — one variable corresponds to toxicity and the other corresponds to efficacy, and one has to consider their joint distribution by incorporating the inherent correlation appropriately. Once a drug is rendered safe from the initial two phases, it is sent up for the Phase III trial for a full-fledged evaluation of its efficacy. The Phase III trial is the most crucial stage where individuals are recruited on a large scale, and investigators often wonder about the number of participants to be recruited so that the effectuality of the drug can be appropriately detected. This is statistically framed as the *problem of sample size determination in clinical trials*. Although, it may appear that the sample size determination problem is encountered only in Phase III trials, it can be faced in the earlier phases as well. However, the methods developed for the Phase III trials can be replicated for Phase I or Phase II. In fact, determining sample sizes is no longer limited to clinical trials and has been actively gaining importance in social, economic, business, and agricultural sciences.

Sample size determination is an important as well as delicate problem, as it determines the cost, time, and efficiency of the trial. A sample size too small may elevate the standard error, whereas an excessive number of participants may increase the time and cost of the

trial without much gain. An anonymous quote reverberates in the research community in this context —

"A small sample speaks in whispers, while a large sample can shout the truth. The key is finding the right volume for your study."

The emergence of the COVID-19 pandemic sparked public interest in vaccines. It was then we found out that even though there is ample literature on sample size about the usual drug trials conducted, the situation was quite opposite as far as the clinical trials of vaccines were concerned. Vaccine trials are particularly different from the ordinary clinical trials, as they aim at estimating the extent of effectiveness of a vaccine in preventing a disease rather than a treatment difference. An elaborate description about vaccine efficacy studies has been given in Sections 1.2 to 1.5 of this chapter.

The problem of determination of sample size will constitute the central premise of this thesis. However, it is to be noted that since sample sizes are calculated at the beginning, investigators often use prior information available at that time. For example, the information available from Phase II or some previously conducted trial may be used to get an idea about the sample size in Phase III. In order to quantify the amount of prior information available, we introduce the notion of Effective Sample Size (ESS), which is closely intertwined with that of sample size determination.

1.1 Effective Sample Size (ESS)

Prior beliefs, also known as co-data, do exist in health care and medical studies, and strategic use of this historical information can lead to better decisions (Neuenschwander et al., 2016; Spiegelhalter et al., 2003). However, the term 'strategic' is the key catch here, as an improper use of the prior information can produce misleading inferences. The Bayesian paradigm has an upper hand over the frequentist approach in this regard, as it can adeptly incorporate the historical information by choosing an appropriate prior distribution on the unknown parameters. The prior distributions are found to be affected by subjectivity on the part of the investigators, and their proper elicitation from the available information is crucial (Savage, 1971; Genest & Zidek, 1986). According to Spiegelhalter et al. (2003, p. 139),

"...turning informally expressed opinions into a mathematical prior distribution is perhaps the most difficult aspect of Bayesian analysis."

Additionally, the prior information might carry some bias, as it could originate from an unreliable source, so the chosen probability distribution should encompass this uncertainty. Therefore, it is of importance to quantify the amount of prior information available

in terms of an ESS (Malec, 2001; Neuenschwander et al., 2010; Pennello & Thompson, 2007). ESSs are defined for a prior distribution to reflect the amount of information contained in the prior — the higher the value of ESS, the more informative the prior is. If the information is dubious, then it is advisable to select a vague prior with low ESS, whereas an informative prior with higher ESS can be selected if the information is more concrete.

ESSs are well understood if the prior chosen on the unknown parameters is conjugate. For example, a binomial model with $Beta(a, b)$ prior has $ESS = a + b$, and the normal prior with variance s_0^2 on normal data with mean μ and known variance σ^2 has ESS $\frac{\sigma^2}{s_0^2}$ (Neuenschwander et al., 2020). ESS can formally be defined in a variety of settings, and the different methods for calculating ESS will be discussed in detail in Section 2.1.1. For now, it is enough to know that most of the calculations will involve computing the prior mean, variances as well as the information matrix. The distributional form of the response variable is equally important as that of the prior. The ESSs will depend on the prior hyper-parameters, and we can define a Posterior ESS by replacing the prior density in the ESS computation formulas by the corresponding posterior density (based on a sample of size n). Neuenschwander et al. (2020) stated a basic *predictive consistency criteria* which can be used to assess the performance of an ESS computation method. It states that for an ideal ESS —

"The expected posterior ESS must be equal to the sum of the prior ESS and the sample size n."

Even though ESS has been studied thoroughly when the responses are univariate and the parameter of interest is single-valued, the literature is scanty when it comes to multi-dimensional parameters and response variables are bivariate (or multivariate). Bivariate responses are often encountered during a Phase II trail, where one response variable reflects the toxicity of the drug and the other reflects efficacy. Modeling their joint distribution by incorporating the inherent correlation appropriately is a primary challenge, especially when one or both of the variables are binary or categorical. In Chapter 2, We concentrate on finding ESS in such situations. The ESSs so obtained can be easily reproduced to a Phase III trial that has bivariate response data.

1.2 Vaccine Efficacy Studies

Vaccine trials, aimed at establishing the benefit of a vaccine, are different from the usual drug (or oncology) trials in various aspects. The differences lie in how "benefit" is perceived, the primary endpoints considered, the individual studied, and the way of trial planning and designing. Vaccine trials emphasize disease prevention, and healthy indi-

viduals are mostly recruited to see the effect of the vaccine on them. Drug trials, on the other hand, focus on treatment responses, and the target population typically consists of patients who are suffering from the disease that the drug is intended to cure. For example, in oncology trials, patients who already have cancer, typically defined by a specific tumor type, stage, and several other factors, are recruited, and changes in their tumor size, progression-free survival, overall survival, etc., post application of the treatment are studied. Another interesting difference is that vaccine trials are often conducted in response to emergency situations (COVID, Ebola) or emerging technologies (mRNA) and often do not fit into the Phase I/II/III schema described at the beginning of the chapter. To speed the development process while maintaining rigor, sponsors may combine two or more phases and call them combined Phase I/II or combined Phase II/III trials. This means that the stages transition within a single protocol under predefined decision rules rather than running separate standalone phases. In May 2020, Pfizer Inc. and BioNTech SE launched the combined Phase I/II trial where healthy volunteers were assigned to different doses ($10\mu g$, $20\mu g$, $30\mu g$, and $100\mu g$) of two candidate RNA-based COVID-19 vaccines, viz., BNT162b1 and BNT162b2. Safety evaluations were done, and ultimately BNT162b2 at $30\mu g$ emerged as the safest dose (Walsh et al., 2020), which then progressed to the combined Phase II/III trial for a full-fledged evaluation of its safety and efficacy. Generally speaking, like every other clinical trial, vaccine trials also proceed from the initial safety screening to the final efficacy confirmation stage. Hence, even though their way of execution may vary, the overall structural pathway of the three phases is similar in both vaccine and drug trials. We follow the regular Phase I, Phase II, and Phase III notation throughout the thesis, but a Phase I may well be replaced by a combined Phase I/II or a Phase III by a combined Phase II/III. The changes in the phase labels are expected to have little to no effect on the methods developed in the thesis.

Safety evaluations in vaccine trials generally look at the occurrences of local and systemic reactions and adverse events due to vaccination. They include mild to severe post-vaccination effects like injection-site pain, swelling, fever, fatigue, headache, etc., to serious life-threatening side effects like hospitalization, death, significant disability, etc. It is quite obvious that a "safe drug" is expected to have such reactions under control. One of the primary safety endpoints in the Pfizer-BioNTech BNT162b2 vaccine trial was solicited local reactions (pain at the injection site, redness, and swelling), systematic events (fever, fatigue, headache, chills, vomiting, diarrhea, new or worsened muscle pain, and new or worsened joint pain), and use of antipyretic or pain medication within 7 days after the receipt of the vaccine or placebo. Apart from this, serious adverse events, unsolicited adverse events, and clinical laboratory abnormalities were also monitored at pre-specified time intervals (Walsh et al., 2020). Similar safety outcomes were also considered in several other COVID-19 vaccine trials and even in the case of Ebola vaccines (Agnandji et al., 2016).

Once the safety check is completed, the Phase III (or combined Phase II/III) trial of a vaccine evaluates its efficacy in comparison to a placebo or other existing vaccine(s) for the disease. The degree of effectiveness of the vaccine in preventing a particular disease is measured by means of protective vaccine efficacy (VE), with a higher value of VE indicating greater efficacy. The mathematical formulation of 'Vaccine efficacy' was first proposed by Greenwood & Yule (1915), which was obtained by comparing the disease-attack rates in the vaccinated and unvaccinated groups.

If ARU and ARV are, respectively, the disease attack rates in the unvaccinated and vaccinated groups, then the ratio of ARV to ARU is called the Risk Ratio (RR). A lower value of RR clearly indicates better performance of the vaccine, since ARV must be lesser than ARU if the vaccine is at all successful in preventing the disease. *One minus RR*, expressed as percentage, is the vaccine efficacy, i.e.

$$VE = \left(1 - \frac{ARV}{ARU}\right) \times 100\% \tag{1.2.1}$$

In other words, VE is nothing but the percentage reduction in the disease-attack rate due to the application of the vaccine, in comparison to placebo. As the *World Health Organization* (WHO) explainers on vaccine development and distribution states,

"If a vaccine has an efficacy of 80 percent: It does not mean that the vaccine will only work 80% of the time. It means that in a vaccinated population, 80% fewer people will contract the disease when they come in contract with the virus."

For example, in December 2020, it was reported that the Pfizer-BioNTech mRNA-based COVID-19 vaccine candidate, BNT162b2 at $30\mu g$, indicated an efficacy rate of 95% during the Phase III trial (Polack et al., 2020), whereas the Oxford AstraZeneca candidate vaccine, ChAdOx1 nCov-2019, appeared to have an overall efficacy of 70.4% during the Phase-III trial (Voysey et al., 2021). Several other studies were conducted to assess the efficacy of the corresponding COVID-19 vaccine candidate. The estimated efficacies of a few of them have been reported in Ghazy et al. (2022).

Although, mathematically VE can take negative values, the scenario is rarely encountered in practice. A negative value indicates that $ARV > ARU$, and it is quite counter-intuitive that a vaccine increases the attack-rate instead of reducing it. Practically speaking, the value of VE ranges between 0% to 100%, with 0% (i.e. $ARV = ARU$) indicating that the vaccine is ineffective, and 100% (i.e. $ARV = 0$) indicating maximum effectiveness.

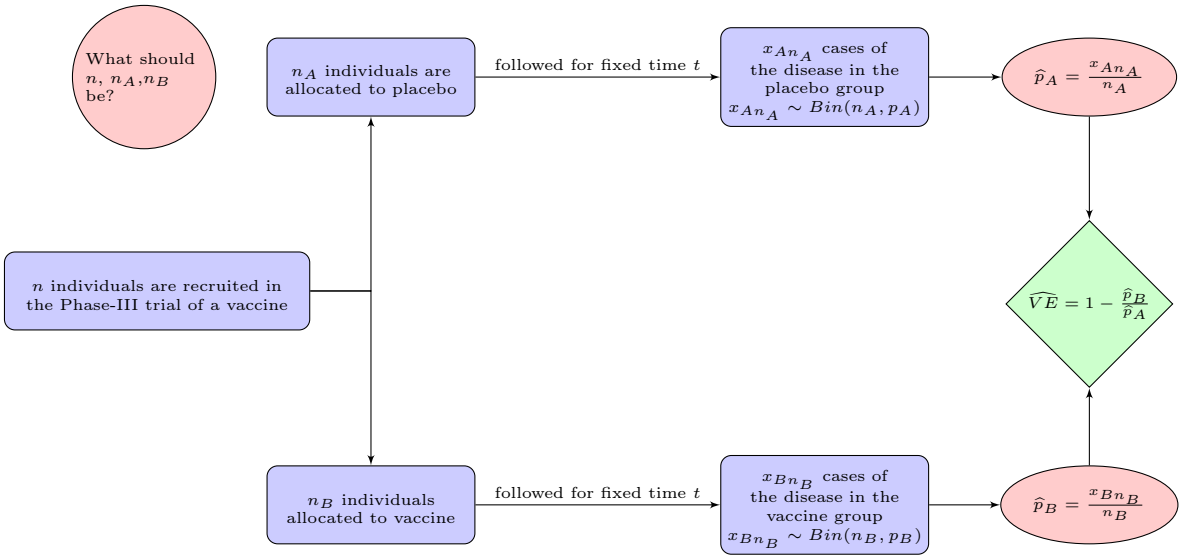

Figure 1.1: A schematic diagram of the sample size determination problem

1.3 Estimation of VE in Comparative Cohort Trials

Cohort studies are prospective designs, and the individuals are followed over time to check for the occurrence of the disease. In Phase III trials, often thousands of individuals are recruited, and they are allocated to either vaccine or placebo. To evaluate the performance of the vaccine, the two cohorts (vaccinated and unvaccinated) are then monitored over several months to see whether the vaccinated cohort get infected at a lower rate than the unvaccinated one (Orenstein et al., 1985, 1988).

Suppose a total of n individuals have been recruited in the trial, of which n_A are allocated to placebo and n_B to vaccine (i.e. $n_A + n_B = n$). Further, let x_{An_A} and x_{Bn_B}, respectively, denote the number of cases of the disease that have occurred after a fixed period of time, in the unvaccinated and vaccinated groups of the cohort.

We assume a binomial model, i.e.

$$\begin{aligned} x_{An_A} &\sim Bin(n_A, p_A), \\ \text{and } x_{Bn_B} &\sim Bin(n_B, p_B), \end{aligned} \tag{1.3.1}$$

independently of each other. Note that ARU and ARV have been renamed as p_A and p_B for the ease of notation. Throughout the thesis, both the notations will be used interchangeably. The estimates of p_A, p_B and subsequently of the Relative Risk Ratio (RR), denoted by ψ are, respectively, $\widehat{p}_A = \frac{x_{An_A}}{n_A}$, $\widehat{p}_B = \frac{x_{Bn_B}}{n_B}$, $\widehat{\psi} = \frac{\widehat{p}_B}{\widehat{p}_A}$. Evidently, VE is estimated as

$$\widehat{VE} = 1 - \widehat{\psi}.$$

Note that model (1.3.1) is assumed to be the underlying model in most vaccine trials and

forms the foundation of Chapters 3 and 4.

It is noteworthy to mention that p_A and p_B may alternatively represent the hospitalization or death rates instead of disease attack rates. This is because vaccine trials may additionally take into account more severe outcomes like hospitalization and death, even though disease occurrence is still the primary outcome. In fact, the BNT162b2 trial reported the efficacy against severe COVID-19 (secondary endpoint), which included criteria such as hospitalization, ICU admission, respiratory failure, or death. The optimistic part is that the methods developed in Chapters 3 and 4 will still be applicable as the meanings of p_A and p_B change, but model (1.3.1) remains unaltered.

The problem of sample size determination generally comprises of finding an appropriate value of n so that a given level of precision (quantified by the length of the CI of VE) or power is achieved in the study. Although it might seem that the determination of the quantity n is of primary interest, finding the best possible split of the total sample size among the vaccine and placebo groups is every bit as significant. In other words, we want to find the ideal values of n_A and n_B, or rather the proportion allocated to placebo, $\rho = \frac{n_A}{n}$. In almost all real-life trials, ρ is fixed at $\frac{1}{2}$ (Equal allocation technique), and approximately an equal number of individuals are allocated across the vaccine and placebo groups. Another popular choice of ρ is $\frac{1}{3}$, known as the Double allocation technique. As the name suggests, here the number of subjects allocated to vaccine is twice that of placebo. However, certain cost and ethical constraints in clinical trials have inspired researchers to deflect from the traditional fixed allocation techniques (with specified ρ) to other allocation schemes where ρ is optimally determined keeping some relevant objectives in mind. For example, the recruitment of each individual in a vaccine trial comes with some added cost, and one should always try to minimize the total sample size without compromising the precision or power. Also, from an ethical point of view, it is advisable to include more individuals to the vaccine group if the vaccine is really effective against preventing the disease. Another reason for randomizing more people to the vaccine group is to increase power for immune correlates analysis (Qin et al., 2007). An immune correlate is a vaccine-induced immune response, measured after vaccination, that is correlated with protective vaccine efficacy. Such correlates are used to predict vaccine efficacy in a new setting, where the estimate of vaccine efficacy may not be directly available. In vaccine trials, analyses are often performed to identify such immune correlates and predict VE based on them. However, immune responses are prone to measurement error, and the existence of such errors can severely reduce the power of the immune correlate analysis. One possible way to reduce such error on immune correlates is to collect more data on them. This is achieved by allocating some additional people to the vaccine group, and one may plan to do a fixed unequal allocation (like Double allocation). But, even in such case, they have to decide how much unequal they want it to be. In other words, the "fixed" proportion of individuals to be assigned to the placebo group (ρ) still needs to be

decided beforehand along with the total number of individuals to be recruited. This again makes it a subclass of the sample size determination problem we are trying to address here. Moreover, the primary objective of hiring more people in the vaccine group is to reduce the measurement error of the immune correlates. Therefore, instead of selecting an ad hoc ρ, it is always advisable to go for one that is optimally determined to give maximal reduction with minimal resource utilization.

1.4 Estimation of VE in Case-Control Studies

Case-control studies are widely used during post-marketing surveillance, after a vaccine has already been licensed for use. Post-marketing surveillance monitors how the vaccine works when it is applied to general masses. Case-control studies are retrospective designs, where a number of cases and controls of the disease are taken, and their vaccination status is checked. The risk ratio is approximated by the Odds Ratio (OR) of cases to controls (Orenstein et al., 1985, 1988), and the vaccine efficacy is estimated as

$$\widehat{VE} = 1 - \widehat{OR}.$$

Suppose n_A controls and n_B cases of the disease are taken ($n_A + n_B = n$), of which y_{An_A} and y_{Bn_B} are, respectively, found to be vaccinated in the two groups. Then, we assume that

$$\begin{aligned} y_{An_A} &\sim Bin(n_A, p_A), \\ \text{and } y_{Bn_B} &\sim Bin(n_B, p_B), \end{aligned} \tag{1.4.1}$$

where p_A and p_B are, respectively, the prevalence of vaccine exposure among controls and cases. The OR, denoted by ξ, is

$$\xi = \frac{p_B(1-p_A)}{(1-p_B)p_A},$$

and it is estimated by

$$\widehat{\xi} = \frac{y_{An_A}/(n_A - y_{An_A})}{y_{Bn_B}/(n_B - y_{Bn_B})}.$$

Odds ratio is the ratio of odds of a person being vaccinated among the cases to that among the controls. If the vaccine provides some resistance against the disease, it is expected that the odds of vaccination among cases should be lower than the odds of being vaccinated in the control group. Subsequently, $\widehat{\xi} < 1$ and $\widehat{VE}$ $(= 1 - \widehat{\xi})$, should be higher for a more effective vaccine.

It is to be noted that even though models (1.3.1) and (1.4.1) look similar, their interpretations are different. But our objective in both cases is to determine the sample size n at the desired level of precision or power. Similar to ρ in the case of cohort trials, determination of the control-to-case ratio $c = \frac{n_A}{n_B}$ is of utmost importance over here. Hence, c may be fixed beforehand by optimizing some criteria (like minimization of the total sample size). In case of fixed allocation, three popular choices of c are $c = 1$ (Equal allocation), $c = 2$ (Double allocation), and $c = 4$ (Quadruple allocation), where the number of controls are respectively equal to, twice, or four times that of the cases.

1.5 Time-Dependent Vaccine Efficacy

As evidenced by Figure 1.1, the participants in the trial, after being assigned to either vaccine or placebo, are monitored for a fixed period of time (say, 2 or 3 months), to examine how many of them have developed the disease by then. It is quite evident that the number of cases of the disease that has been observed will vary according to the time to follow up. Thus, in reality, vaccine efficacy will depend on time t, and equation (1.2.1) should be modified as

$$VE(t) = 1 - \frac{ARV(t)}{ARU(t)}. \tag{1.5.1}$$

Note that, the disease attack rates in the vaccinated and unvaccinated populations at time t, respectively denoted by $ARV(t)$ and $ARU(t)$, are actually the probabilities that a person in the respective populations will contract the disease within a time-period t subsequent to recruitment.

The estimates of vaccine efficacy obtained at two different time-points often differ. The final estimate of VE reported from a trial, apparently calculated using formula (1.2.1), is actually the estimate taken at the end of the follow-up period. In other words, estimate of $VE(L)$ is reported, where L is the length of the follow-up period of each individual in the study. If, for two different time-points t_1 and t_2 $(t_1 < t_2)$, the data provide evidence that $VE(t_1) > VE(t_2)$, then investigators may conclude that the effect of the vaccine wanes over time. In fact, several analyses have provided evidence that the various COVID-19 vaccines have a waning effect (Ferdinands et al., 2022; Feikin et al., 2022). Feikin et al. (2022) performed an analysis to evaluate the duration of effectiveness of vaccines against SARS-Cov-2 infection and COVID-19 disease. It included 78 vaccine-specific vaccine efficacy evaluations of 5 vaccines viz. Pfizer–BioNTech-Comirnaty, Moderna-mRNA-1273, Janssen-Ad26.COV2.S, and AstraZeneca-Vaxzevria. On an average, it was seen that the overall vaccine efficacy against SARS-CoV-2 infection decreased from 1 month to 6 months after full vaccination by around 21%. The decline was more pronounced in case of symptomatic COVID-19 disease than severe COVID-19 disease. However, to conclude that the effect of the vaccine fades over time based is statistically incorrect as both $ARV(t)$

and $ARU(t)$ increases with t, and a simultaneous increase of both may lead to a decrease in the value of $VE(t)$. But, susceptibility to the disease in the vaccinated population as compared to the unvaccinated population may remain unchanged with time. Therefore, so far time-dependent vaccine efficacy (VE) is concerned, its definition in terms of relative risk (Orenstein et al., 1985) is preferred, which states that

$$VE(t) = 1 - RR(t), \tag{1.5.2}$$

where $RR(t)$ is the relative risk (at time t) of getting affected in the vaccinated group as compared to the unvaccinated group.

If one considers VE to be time-invariant and assumes a binomial model in estimating VE (O'Neill, 1988; Bose & Biswas, 2023), the two definitions coincide since $\frac{ARV}{ARU}$ indeed becomes the relative risk. Although the binomial model is widely used in the trials, it may not be appropriate as the number of cases of the disease developed increase with the passage of time. A survival model by considering the time-to-event (time to development of the disease or time to hopspitalization/death if we are considering efficacy against severe disease) data would be more appropriate in such a scenario. Using this approach, we can test or estimate the waning effect of the vaccine (Durham et al., 1998).

1.6 Structural Organization of the Thesis

This thesis illuminates the problem of sample size determination in the context of vaccine efficacy trials from various perspectives — both frequentist and Bayesian. The problem of finding ESS has also been explored initially. An overview of each chapter including the major findings from them is summarized as follows.

1.6.1 Overview of Chapter 2

Neuenschwander et al. (2020) reviewed the existing methods for calculating ESS in a clinical trials set up. They reviewed five methods, viz. variance-ratio (VR), precision-ratio (PR), Morita-Thall-Müller (MTM), a modified version of Morita-Thall-Müller (MTM.P), and Expected Local Information Ratio (ELIR) for calculating ESS. They introduced the Expected Local Information Ratio (ELIR) method and showed that it satisfies the basic predictive consistency criterion. However, they considered only the univariate case as the variable corresponding to the efficacy measure was assumed to be the only component of the treatment-response vector. In a discussion, Biswas & Angers (2020) extended the approach of finding ESS in the bivariate case where they considered efficacy as well as toxicity as two components of the treatment-response vector. However, both the efficacy

and safety measures were considered to be binary, which may not be the case in a typical phase II trial. In Chapter 2, we extend the approach further and consider the situation when one of the components of the response vector is binary and another is continuous. Without loss of generality, we denote Y to be the (binary) safety measure, and Z the (continuous) efficacy measure, and the following joint model (Cheng & Biswas, 2008) is considered for the bivariate response variable (Y, Z) :

$$P(Y = 1) = p,\ 0 < p < 1 \text{ and } Z|Y = y \sim N(\mu_y, \sigma^2),\ y = 0, 1. \tag{1.6.1}$$

The common variance σ^2 is initially assumed to be known. Two types of prior are assumed on the parameter $\theta = (p, \mu_0, \mu_1)'$, of which one is a Beta and Bivariate Normal prior on p and $\boldsymbol{\mu} = (\mu_0, \mu_1)'$, respectively. In the other case, a Multivariate-T distribution is taken on $\boldsymbol{\mu}$ instead of Bivariate Normal. The ESSs for these priors were computed using the VR, PR, MTM, MTM.P and ELIR methods and the results were compared among themselves. It was seen that all the methods provided similar results when the prior was more informative. Similar observations were made when we considered σ^2 to be unknown and took an additional Inverse Gamma prior on it. The methods for calculating ESS using all the five methods were outlined, and the results were compared among themselves. It was observed that in almost all the cases the ESS values calculated using different methods became more and more similar as we moved towards a more informative prior from a vague prior.

It was already clear from the expressions that the ESSs did not satisfy the *the predictive consistency criteria*. The best we could do was to compare the expected posterior ESS, subtracted by n, with the prior ESS for different values of n and observe its deviation from the predictive consistency criteria. The Average Posterior ESS were computed by simulation, and it was seen that the performance of the methods were equally poor when it came to evaluating the deviation from the predictive consistency criteria. In the next part of the study, we introduced a a double-blind, randomized, multi-centre, phase II trial reported in Ella et al. (2021), that had a binary safety and continuous efficacy data. The information from this data has been used as historical information, and the hyper-parameters of a prior was selected using this information. The ESS values using the different methods were calculated, and their relative performances were evaluated. Overall, it was seen that VR and ELIR methods provided better results even when the priors were vague.

1.6.2 Overview of Chapter 3

As it is already mentioned, not much work have been done on how to determine sample size in vaccine trials. In one of the early works by O'Neill (1988), the problem in both

comparative cohort and case control designs was approached by presuming models (1.3.1) and (1.4.1). In *comparative cohort designs*, a $100(1-\alpha)\%$ CI for $\beta = \log \psi$ was obtained using the large sample properties of binomial distribution. The desired width of the CIs was pre-specified, and the minimum sample size required to achieve that width was calculated. But it was assumed that an equal number of individuals were allocated across the vaccine and placebo groups. In Chapter 3, we build on this concept and introduce methods for calculating the required sample size with a specified degree of precision when there is an unequal allocation of individuals across the two groups. The allocation proportion ρ was chosen in such a manner that the total sample size or the expected number of people developing the disease is minimized, and they respectively gave rise to Neyman and RSIHR-type allocations. The optimum allocations were called so due to their analogies with the Neyman (Melfi & Page, 1998) and RSIHR allocations (Rosenberger et al., 2001). The optimum ρs were derived as,

$$\rho_{Neyman} = \frac{\sqrt{(1-p_A)p_B}}{\sqrt{(1-p_A)p_B} + \sqrt{(1-p_B)p_A}}, \text{ and } \rho_{RSIHR} = \frac{\sqrt{1-p_A}p_B}{\sqrt{1-p_A}p_B + \sqrt{1-p_B}p_A}.$$

The total sample size and the sample size required for the vaccine group under the different allocation schemes (Equal, Double, Neyman and RSIHR) were computed by setting different values of the desired width (or relative width with respect to VE), VE and ARU. It was found that Neyman allocation provided the lowest sample sizes, while RSIHR allocated higher number of individuals to the vaccine group for higher vaccine efficacy. Overall, it could be concluded that it is better to go for Neyman or RSIHR allocation if reducing the sample size or the average number of people developing the disease is of interest.

At times, there may be a limited vaccine supply and the number of people that can be allocated to vaccine will be restricted. In such scenarios, the optimum allocation ratios need to be slightly modified to compensate for the same. The method for obtaining modified versions of the Neyman and RSIHR allocations under restricted vaccine dosage was also outlined. Another approach was followed, where we obtained the optimal sample sizes by fixing the power of the relevant test for VE, instead of the width/relative width of the CI for VE.

For *case-control* studies, a similar approach was followed to obtain the required sample sizes and optimal control-to-case ratio. But here, the confidence interval (CI) for VE was derived using the asymptotic distribution of $\widehat{\phi} = \log \widehat{\xi}$. The conclusions were similar to that of the comparative cohort trials. In the next part, we compared our sample sizes to the actual sample sizes taken in a real-life randomized clinical trial. It was seen that, under some assumptions, if Neyman allocation was used instead of the actual (1:1) allocation, the same relative width would have been achieved by a much lesser number of

samples. RSIHR allocation provided nearly equal sample sizes as the actual one, but it allocated more individuals to the vaccine group, thereby reducing the overall chance of a person developing the disease. It should also be noted that the numerical values of vaccine efficacy (VE) and the disease attack-rate in the general population (ARU) are unknown, and are to be estimated from the trial. Hence, at the beginning of the trial, to get an idea about the required sample size, we try to guess the unknown values VE and ARU using some prior information, usually available from some previously conducted trial. If the individuals are being recruited sequentially, we do not know when to stop recruiting or how to allocate the incoming individuals to the vaccine or placebo group so that our desired allocation proportion (ρ_{Neyman} or ρ_{RSIHR}) is achieved. So, the implementations of the allocations in real-life by using Response Adaptive Randomization (RAR) techniques like Sequential Maximum Likelihood Estimation (SMLE) (Rosenberger et al., 2001; Zhang et al., 2006; Biswas et al., 2007; Ivanova & Rosenberger, 2001), Doubly Biased Coin Design (DBCD) (Eisele, 1994; Eisele & Woodroofe, 1995; Hu & Zhang, 2004; Melfi et al., 2001) and Effective Randomized-Adaptive Designs (ERADE) (Hu et al., 2009) procedures were also shown.

1.6.3 Overview of Chapter 4

The RAR techniques suffered from certain drawbacks, and even though they helped us achieve the desired allocation asymptotically, we still had to rely on the guesses of VE and ARU to get an idea about the initial sample size. A wrong or misleading assumption regarding these values will give disastrous results. This flaw can be countered if we follow a Bayesian approach as it can deal with the uncertainities associated with the unknown values of the parameters (Adcock, 1988). Hence, in Chapter 4, we propose a Bayesian alternative to the frequentist sample size determination problem and highlight the advantage of the former over the latter. The same model (1.3.1) was considered and independent $Beta(a_A, b_A)$ and $Beta(a_B, b_B)$ priors were considered on p_A and p_B. Instead of the CI, we obtained the Highest Posterior Density (HPD) intervals for VE, whose length and coverage were obtained using two approaches: one by applying Monte Carlo techniques and another by using the asymptotic posterior distribution of VE. The usual criterion functions, like the Average Coverage Criterion (ACC) and Average Length Criterion (ALC) (details to be discussed later on) used in Bayesian sample size determination, were deployed for the calculation of the sample sizes. In fact, optimal allocations analogous to that Neyman and RSIHR were obtained here as well. The sample sizes as well as the optimal allocation proportions came out to be a function of the hyper-parameters $\boldsymbol{\eta} = (a_A, b_A, a_B, b_B)'$, and their values were chosen in such a way that their mean reflect the assumed values of ARU and ARV, and variances deal with the uncertainity associated with them. But one drawback of the Bayesian approach was that sample sizes were

not easy to calculate and there were no closed form formulas as the frequentist version. Hence, we derived certain bounds on the sample sizes that were easy to calculate, and somewhat solved the computational issues. The entire procedure was then replicated for case-control studies.

Both the Bayesian and frequentist sample sizes were then compared to the sample sizes taken in a real-life Phase III trial. The Bayesian sample sizes reduced as the prior became more informative. The sample sizes were even lesser than the frequentist sample sizes for a slightly informative prior. The Bayesian method was equally successful as its frequentist counterpart in achieving the desired objective when one deflected from the conventional equal (or fixed) allocation technique to other optimal allocation schemes (Neyman or RSIHR). Overall, we could say that the Bayesian sample sizes are preferred over the frequentist one as they simultaneously prevents the loss of precision even if there are uncertainities about the underlying values of parameters, and fulfills the optimal allocation objectives. However, the entire procedure depends on the proper elicitation of prior.

1.6.4 Overview of Chapter 5

We discussed in Section 1.5 that the usual binomial model described in section 1.2 may fail to capture the dynamic nature of the vaccine trial. A survival approach that considers the time-to-event (time to development of the disease) data can provide better results. Furthermore, the survival approach elegantly handles the presence of censoring in the data in the form of loss-to-follow up or drop out from the study. The binomial model, on the other hand, cannot handle such issues and introduces bias in the estimates of ARV and ARU. In order to adjust for the differences in time-to-follow-up for various individuals, investigators often report the disease attack rates in terms of per 1000 person years of follow-up. But a survival method would be statistically more appropriate and obviously would provide better results. However, it is to be noted that both Chapters 3 and 4 considered the basic binomial model and ignored the time-to-event nature of the vaccine trials.

In the survival framework, the definition of vaccine efficacy in terms of $RR(t)$ is considered, which is nothing but the hazard ratio. The $VE(t)$, given by equation (1.5.2), then reduces to

$$VE_H(t) = 1 - \frac{\lambda_B(t)}{\lambda_A(t)}, \tag{1.6.2}$$

where $\lambda_A(t)$ and $\lambda_B(t)$, respectively, denote the hazard rate of the unvaccinated and vaccinated populations at time t (Durham et al., 1998). To avoid confusion with the former definition of vaccine efficacy (equation (1.5.1)), we denote the vaccine efficacy

defined in terms of the hazard ratio by $VE_H(t)$. Suppose

$$\gamma(t) = \log \frac{\lambda_B(t)}{\lambda_A(t)}. \tag{1.6.3}$$

Thus, $VE_H(t)(= 1 - e^{\gamma(t)})$ remain independent of time if $\gamma(t) = \gamma$ for all t. However, $VE(t)$ (equation (1.5.1)) may still depend on time; and its estimate may provide a false sense of waning discussed earlier. $VE_H(t)$ comes to the rescue in such a case and a test for

$$H_0 : \gamma(t) = \gamma \text{ for all } t$$

must be performed to check whether vaccine efficacy is really time-dependent or not.

It is quite obvious that a change in the underlying model will affect the sample size calculations, and the problem needs to be restructured in the light of the new model. In Chapter 5, we recognize the time-to-event nature of the vaccine trial data and approach the sample size determination problem in a survival analysis framework. We observe that the survival method outperforms the traditional binomial model. We assume $VE_H(t)$ to be independent of t and obtain a $100(1-\alpha)\%$ CI of its common value VE_H, so that the total sample size can be determined.

If T_A and T_B denote the survival times in the respective placebo and groups, we assume two distributions, viz., exponential with rate λ_J and Weibull with parameters λ_J and p, on T_J, $J = A, B$. A semi-parametric approach, which do not require any distributional assumptions on T_A and T_B, and used the Cox Proportional Hazards model (Cox, 1972) in estimating VE_H, was also considered. In fact, when the actual survival times followed exponential or Weibull distribution, the outcomes from the semi-parametric method were almost identical to those obtained from the standard parametric method.

When T_J's are assumed to be $Exp(\lambda_J)$, $J = A, B$, the sample sizes came out to be a function of λ_A, VE_H (or λ_B), coverage probability α, desired width W and the length of the follow-up period L. Neyman and RSIHR-like allocations which minimize the total sample size or the expected number of disease occurrences within the study period were developed, and they came out to be a function of λ_A, VE_H and L. The sample sizes decreased drastically with the decrease in λ_A or increase in VE_H, provided all other factors are kept constant. Additionally, the sample sizes were comparatively higher when the follow-up time was smaller or the amount of censoring was large to compensate for the loss of information that a shorter study or higher censoring may incur. Similar observations were made when we considered Weibull survival times, or took the semi-parametric approach.

Apart from this, the other aspects of the sample size determination problem like obtaining sample sizes by fixing the power of the relevant test for VE_H, or the methods

for achieving the optimal allocation proportions using RAR techniques (SMLE, DBCD or ERADE) were also shown. All the sample sizes, including the ones obtained by considering the traditional Binomal model were compared to the actual sample sizes taken in a real-life trial. The survival approach clearly outperformed the binomial model, and the Neyman and RSIHR allocations fulfilled their respective objectives.

1.6.5 Overview of Chapter 6

Despite thoroughly examining the unexplored area of sample size determination in vaccine trials, there are certain avenues that need further excavation. In Chapter 6, we shed light on some of these unexplored areas and point out possible future research directions before finally drawing the curtain.

Chapter 2

Effective Sample Size in Phase II Trials

2.1 Introduction

It has already been mentioned in Chapter 1 that ESS are defined in the context of Bayesian Analysis in order to quantify the amount of information the prior contains. In this chapter, we consider the problem of finding ESS in Phase II clinical trials where toxicity and efficacy are the two components of the treatment response vector. In particular, one of the components is assumed to be binary and the other is assumed to be continuous. The case of binary safety and continuous efficacy is studied for different prior distributions (both conjugate and non-conjugate) under various scenarios.

Neuenschwander et al. (2020) reviewed the existing methods for calculating ESS which can be applied to non-conjugate families as well. The five methods viz. variance-ratio (VR), precision-ratio (PR), Morita-Thall-Müller (MTM), a modified version of Morita-Thall-Müller (MTM.P), and Expected Local Information Ratio (ELIR) discussed by them are briefly described below.

2.1.1 Different Methods for Obtaining ESS

Suppose we have a sample of size n, viz. $\{Y_1, Y_2, \ldots, Y_n\}$, on a univariate response variable Y with density $f(y|\theta)$. If $\widehat{\theta}_n = \widehat{\theta}_n(Y_1, Y_2, \ldots, Y_n)$ is an estimator of θ based on the sample, then one possible way of calculating ESS is by making use of the prior variance of θ and the (expected) variance of $\widehat{\theta}_n$. The ESS is defined as the n for which the two variances are the same (Neuenschwander et al., 2020) and its expression is given by,

$$ESS = n\frac{E_\theta(Var(\hat{\theta}_n))}{Var(\theta)}. \tag{2.1.1}$$

The above-defined ESS serve as the basis when comparing ESS values from the different methods and has often been considered as the actual ESS.

The variance-ratio (VR) and the precision-ratio (PR) methods make use of the prior variance and the Fisher's information matrix for calculating the ESSs. The expressions are given by

$$ESS_{VR} = \frac{E_\theta\{i_F^{-1}(\theta)\}}{Var(\theta)},\ ESS_{PR} = \frac{Var^{-1}(\theta)}{E_\theta\{i_F(\theta)\}},$$

where $i_F(\theta) = E_{Y_1|\theta}\{i_F(Y_1;\theta)\} = -E_{Y_1|\theta}\left\{\frac{d^2 \log f(Y_1|\theta)}{d\theta^2}\right\}$ is the expected Fisher information based on unit observation.

Morita et al. (2008) suggested an ESS by minimizing the distance (evaluated at the prior mean $\bar{\theta}$) between the expected posterior information obtained from a sample of size n by taking a vague prior $p_0(\theta)$ with the same mean $\bar{\theta}$ and the information from the actual prior $p(\theta)$. This method is called the MTM method and its expression for ESS is given by

$$ESS_{MTM} = \frac{i(p(\bar{\theta})) - i(p_0(\bar{\theta}))}{E_{Y_1}\{i_F(Y_1;\bar{\theta})\}},$$

where $i(p(\theta)) = -\frac{d^2 \log p(\theta)}{d\theta^2}$ and $i(p_0(\theta)) = -\frac{d^2 \log p_0(\theta)}{d\theta^2}$ are, respectively, the information from the prior $p(\theta)$ and the vague prior $p_0(\theta)$. Here $E_{Y_1}\{i_F(Y_1;\theta)\}$ is the expectation of the observed information matrix $i_F(Y_1;\theta)$, where the expectation is taken over the prior-predictive distribution.

The information from the vague prior is negligible and the term $i(p_0(\bar{\theta}))$ can be ignored. Another form of ESS is obtained in somewhat similar fashion by taking the prior mode $(\widetilde{\theta})$ instead of the prior mean $\bar{\theta}$, ignoring the term $i(p_0(\bar{\theta}))$ and approximating $E_{Y_1}\{i_F(Y_1;\theta)\}$ by $i_F(\theta)$. It is defined as

$$ESS_{MTM.P} = \frac{i(p(\widetilde{\theta}))}{i_F(\widetilde{\theta})}.$$

Neuenschwander et al. (2020) even introduced a new method, viz. Expected Local Information Ratio (ELIR) that calculates ESS by computing the expectation (over θ) of the prior information to the Fisher information ratio $r(\theta)$, i.e.

$$ESS_{ELIR} = E_\theta(r(\theta)) = E_\theta\left\{\frac{i(p(\theta))}{i_F(\theta)}\right\}.$$

The above ESSs have been computed using the prior density of θ, and by 'ESS' we usually refer to this Prior ESS. Neuenschwander et al. (2020) proposed the *predictive consistency criteria*, which should be satisfied by an ideal ESS. However, only ESS_{ELIR} was shown to be predictively consistent. Nevertheless, the rest of the ESSs satisfied the criteria asymptotically as the difference between the Expected Posterior ESS and n

converged towards the corresponding Prior ESS when n was sufficiently large. But the aforementioned ESSs were defined using an univariate response variable and a single-valued parameter of interest. The examples in Section 3 of Neuenschwander et al. (2020) considered randomized proof-of-concept phase II trials where they considered the variable corresponding to the efficacy measure to be the only component of the treatment-response vector. In a subsequent discussion paper, Biswas & Angers (2020) extended the approach of finding ESS in the bivariate case, where they considered efficacy as well as toxicity as two components of the treatment-response vector. However, both the efficacy and safety measures were considered to be binary, which may not be the case in a typical phase II trial.

2.1.2 Set Up

We extend the approach further and consider the situation when one of the components of the response vector is binary and another is continuous. Without loss of generality, we consider the safety measure to be binary (typically, non-toxic or toxic, interpreted as success or failure, respectively) and the efficacy measure to be continuous.

Let Y denotes the (binary) safety measure ($Y = 1$ or 0 according to a success or failure), and Z denotes the (continuous) efficacy measure (say, a higher value of Z indicating a more favourable response). Clearly, (Y, Z) being the safety and efficacy measures of the same individual, they are (possibly) correlated, and we need a joint probability model for (Y, Z).

To model the mixed response variables (Y, Z), we consider Y to be a Bernoulli random variable with success probability p; and conditionally given Y, the efficacy measure Z is assumed to be normally distributed. The mathematical formulation of the model is given in equation (1.6.1). Initially, the common variance σ^2 is assumed to be known but the presumption is changed later on.

The rest of the chapter is organized as follows. In Section 2.2, we consider independent Beta prior for p and a Bivariate normal prior for $(\mu_0, \mu_1)'$. The priors are conjugate in this case. In Section 2.3, we obtain the ESS values for the prior considered in Section 2.2. In Section 2.4, we assume a Multivariate-T prior instead of a Bivariate Normal prior for $(\mu_0, \mu_1)'$ and obtain the different ESSs. The remaining assumptions are the same as in Section 2.2. In Section 2.5, we assume the common variance σ^2 to be unknown, and assume an additional inverse gamma prior for σ^2 along with the independent Beta and Bivariate Normal prior considered in Section 2.2. The expressions for ESS are updated in such a case. In Section 2.6, we compare the prior ESS and posterior ESS for different values of the hyper-parameters of the priors considered in Section 2.2, 2.4 and 2.5. In Section 2.7, we chose our prior using the information from a real-life dataset and obtain

the corresponding ESSs. Section 2.8 concludes.

2.2 Mixed Response Model with Independent Beta and Bivariate Normal Prior

The value of σ^2 is assumed to be known in this section. The joint density of (Y, Z) is given by

$$\begin{aligned} f(y, z|\boldsymbol{\theta}) &= f_1(y|p).f_2(z|y, \boldsymbol{\mu}, \sigma^2) \\ &= p^y(1-p)^{1-y}.\frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(z - \mathbf{x}'\boldsymbol{\mu})^2\right\}, \end{aligned}$$

where

$$\mathbf{x} = \begin{pmatrix} 1-y \\ y \end{pmatrix}, \ \boldsymbol{\mu} = \begin{pmatrix} \mu_0 \\ \mu_1 \end{pmatrix} \text{ and } \boldsymbol{\theta} = \begin{pmatrix} p \\ \mu_0 \\ \mu_1 \end{pmatrix}.$$

The Fisher Information matrix based on a single observation (Y, Z) is given by,

$$i_F(\boldsymbol{\theta}) = E_{(Y,Z)|\boldsymbol{\theta}}[i_F\{(Y, Z); \boldsymbol{\theta}\}] = \begin{bmatrix} \frac{1}{p(1-p)} & 0 & 0 \\ 0 & \frac{1-p}{\sigma^2} & 0 \\ 0 & 0 & \frac{p}{\sigma^2} \end{bmatrix}.$$

For n observation vectors on (Y, Z), given by $\{(y_1, z_1), (y_2, z_2), \ldots, (y_n, z_n)\}$, the likelihood is

$$\begin{aligned} L(\boldsymbol{\theta}|(y_i, z_i), i = 1, 2, \ldots, n) &= \prod_{i=1}^{n}\{f_1(y_i|p).f_2(z_i|y_i, \boldsymbol{\mu})\} \\ &\propto p^{n_1}(1-p)^{n_0}.\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(z_i - \boldsymbol{x_i'}\boldsymbol{\mu})^2\right\}, \end{aligned}$$

where $n_1 = \sum_{i=1}^{n} y_i$ and $n_0 = n - n_1$. We take independent priors for p and $\boldsymbol{\mu}$, i.e. we assume

$$\begin{aligned} &p \sim Beta(a, b), \ a > 0, b > 0 \text{ and} \\ &\boldsymbol{\mu} \sim N_2(\boldsymbol{\gamma}, \Sigma), \text{ independently of each other,} \end{aligned} \tag{2.2.1}$$

where $\boldsymbol{\gamma} = \begin{pmatrix} \gamma_0 \\ \gamma_1 \end{pmatrix}$ and $\Sigma = \begin{bmatrix} \sigma_0^2 & \rho\sigma_0\sigma_1 \\ \rho\sigma_0\sigma_1 & \sigma_1^2 \end{bmatrix}$, $\gamma_0, \gamma_1 \in \mathbb{R}$ and $\sigma_0^2, \sigma_1^2 > 0, \ 0 < \rho < 1$.

The posterior of $\boldsymbol{\theta}$ calculated from the prior and the likelihood, and is given by

$$\begin{aligned} &p|\{(y_i, z_i), i = 1, 2, \ldots, n\} \sim Beta(a + n_1, b + n_0), \text{ and} \\ &\boldsymbol{\mu}|\{(y_i, z_i), i = 1, 2, \ldots, n\} \sim N_2(\boldsymbol{\gamma}^*, \Sigma^*), \text{ independently of each other,} \end{aligned} \tag{2.2.2}$$

where $\boldsymbol{\gamma}^* = \begin{pmatrix} \gamma_0^* \\ \gamma_1^* \end{pmatrix}$ and $\Sigma^* = \begin{bmatrix} \sigma_0^{*^2} & \rho^*\sigma_0^*\sigma_1^* \\ \rho^*\sigma_0^*\sigma_1^* & \sigma_1^{*^2} \end{bmatrix}$, with

$$\gamma_0^* = \frac{\left\{\left(\frac{n_1}{\sigma^2} + \frac{1}{\sigma_1^2(1-\rho^2)}\right)\left(\frac{\sum_{i=1}^n z_i(1-y_i)}{\sigma^2} + \frac{\gamma_0}{\sigma_0^2(1-\rho^2)} - \frac{\rho\gamma_1}{\sigma_0\sigma_1(1-\rho^2)}\right) + \frac{\rho}{\sigma_0\sigma_1(1-\rho^2)}\left(\frac{\sum_{i=1}^n z_i y_i}{\sigma^2} + \frac{\gamma_1}{\sigma_0^2(1-\rho^2)} - \frac{\rho\gamma_0}{\sigma_0\sigma_1(1-\rho^2)}\right)\right\}}{\left(\frac{n_0}{\sigma^2} + \frac{1}{\sigma_0^2(1-\rho^2)}\right)\left(\frac{n_1}{\sigma^2} + \frac{1}{\sigma_1^2(1-\rho^2)}\right) - \frac{\rho^2}{\sigma_0^2\sigma_1^2(1-\rho^2)^2}},$$

$$\gamma_1^* = \frac{\left\{\left(\frac{n_0}{\sigma^2} + \frac{1}{\sigma_0^2(1-\rho^2)}\right)\left(\frac{\sum_{i=1}^n z_i y_i}{\sigma^2} + \frac{\gamma_1}{\sigma_1^2(1-\rho^2)} - \frac{\rho\gamma_0}{\sigma_0\sigma_1(1-\rho^2)}\right) + \frac{\rho}{\sigma_0\sigma_1(1-\rho^2)}\left(\frac{\sum_{i=1}^n z_i(1-y_i)}{\sigma^2} + \frac{\gamma_0}{\sigma_1^2(1-\rho^2)} - \frac{\rho\gamma_1}{\sigma_0\sigma_1(1-\rho^2)}\right)\right\}}{\left(\frac{n_0}{\sigma^2} + \frac{1}{\sigma_0^2(1-\rho^2)}\right)\left(\frac{n_1}{\sigma^2} + \frac{1}{\sigma_1^2(1-\rho^2)}\right) - \frac{\rho^2}{\sigma_0^2\sigma_1^2(1-\rho^2)^2}},$$

$$\sigma_0^{*^2} = \frac{\frac{n_1}{\sigma^2} + \frac{1}{\sigma_1^2(1-\rho^2)}}{\left(\frac{n_0}{\sigma^2} + \frac{1}{\sigma_0^2(1-\rho^2)}\right)\left(\frac{n_1}{\sigma^2} + \frac{1}{\sigma_1^2(1-\rho^2)}\right) - \frac{\rho^2}{\sigma_0^2\sigma_1^2(1-\rho^2)^2}},$$

$$\sigma_1^{*^2} = \frac{\frac{n_0}{\sigma^2} + \frac{1}{\sigma_0^2(1-\rho^2)}}{\left(\frac{n_0}{\sigma^2} + \frac{1}{\sigma_0^2(1-\rho^2)}\right)\left(\frac{n_1}{\sigma^2} + \frac{1}{\sigma_1^2(1-\rho^2)}\right) - \frac{\rho^2}{\sigma_0^2\sigma_1^2(1-\rho^2)^2}},$$

$$\text{and } \rho^* = \frac{\rho}{\sqrt{\left(1 + \frac{\sigma_0^2}{\sigma^2}(1-\rho^2)n_0\right)\left(1 + \frac{\sigma_1^2}{\sigma^2}(1-\rho^2)n_1\right)}}.$$

The derivation of the posterior is shown in Appendix A.1 and it is quite clear that the prior of $\boldsymbol{\theta}$ is conjugate in this case.

2.3 Expressions for ESS

Note that, an expression for the univariate ESS is given by (2.1.1).

As in Biswas & Angers (2020), we also considered a scalarized version of (2.1.1) to obtain the ESS for our multivariate model. Even though both the trace and determinant versions can be considered, we only report the determinant version for the sake of brevity.

Thus, the ESS is given by

$$ESS = n\{|Var(\theta)|^{-1}|E(Var(\hat{\theta}_n))|\}^{\frac{1}{3}}. \tag{2.3.1}$$

We will consider this to be the basic expression for ESS and compare the different methods against it.

For model (1.6.1) with known σ^2, the maximum likelihood estimate (MLE) of θ is given by $\hat{\theta}_n = (\hat{p}_n, \hat{\mu}_{0n}, \hat{\mu}_{1n})'$, where

$$\begin{aligned}\hat{p}_n &= \frac{1}{n}\sum_{i=1}^{n} y_i,\\ \hat{\mu}_{0n} &= \frac{\sum_{i=1}^{n}(1-y_i)z_i}{\sum_{i=1}^{n}(1-y_i)} \text{ and}\\ \hat{\mu}_{1n} &= \frac{\sum_{i=1}^{n} y_i z_i}{\sum_{i=1}^{n} y_i}.\end{aligned}$$

Using the above estimate and also (2.3.1), we obtain the ESS of prior (2.2.1) as

$$ESS = \left\{\frac{\sigma^4}{\sigma_0^2\sigma_1^2(1-\rho^2)}.\frac{(a+b)(a+b-1)^2}{(a-1)(b-1)}\right\}^{\frac{1}{3}}, \text{ provided } a>1, b>1. \tag{2.3.2}$$

Similarly, the expressions for ESSs are calculated using the different methods mentioned in Section 2.1.1. The expressions for ESSs along with their formulae are shown in the Table 2.1. We see that VR and $ELIR$ provides values exactly equal to the actual ESS

ESS	**Formula**	**Value**
ESS_{VR}	$\left\{\lvert Var(\theta)\rvert^{-1}\lvert E\{i_F^{-1}(\theta)\}\rvert\right\}^{\frac{1}{3}}$	$\left\{\frac{\sigma^4}{\sigma_0^2\sigma_1^2(1-\rho^2)}.\frac{(a+b)(a+b-1)^2}{(a-1)(b-1)}\right\}^{\frac{1}{3}}$
ESS_{PR}	$\{\lvert E\{i_F(\theta)\}\rvert^{-1}\lvert Var^{-1}(\theta)\rvert\}^{\frac{1}{3}}$	$\left\{\frac{\sigma^4}{\sigma_0^2\sigma_1^2(1-\rho^2)}.\frac{(a+b)^4(a+b+1)(a-1)(b-1)}{a^2b^2(a+b-1)(a+b-2)}\right\}^{\frac{1}{3}}$
ESS_{MTM}	$\left\{\lvert E_{(Y,Z)}\{i_F((Y,Z);\overline{\theta}))\}\rvert^{-1}\lvert i(p(\overline{\theta}))\rvert\right\}^{\frac{1}{3}}$	$\left\{\frac{\sigma^4}{\sigma_0^2\sigma_1^2(1-\rho^2)}.(a+b)^2\left(\frac{a-1}{a^2}+\frac{b-1}{b^2}\right)\right\}^{\frac{1}{3}}$
$ESS_{MTM.P}$	$\left\{\lvert E\{i_F(\widetilde{\theta}))\}\rvert^{-1}\lvert i(p(\widetilde{\theta}))\rvert\right\}^{\frac{1}{3}}$	$\left\{\frac{\sigma^4}{\sigma_0^2\sigma_1^2(1-\rho^2)}.\frac{(a+b-2)^3}{(a-1)(b-1)}\right\}^{\frac{1}{3}}$
ESS_{ELIR}	$\left\{\lvert E_\theta\{i_F^{-1}(\theta)i(p(\theta))\}\rvert\right\}$	$\left\{\frac{\sigma^4}{\sigma_0^2\sigma_1^2(1-\rho^2)}.\frac{(a+b)(a+b-1)^2}{(a-1)(b-1)}\right\}^{\frac{1}{3}}$

Table 2.1: Formulae for calculating Multivariate ESS along with their values for model (1.6.1) and prior (2.2.1).

given by (2.3.2). Also, it can be easily seen that all values of ESS are similar and approximately equal to

$$\left\{\frac{\sigma^4}{\sigma_0^2\sigma_1^2(1-\rho^2)}.\frac{(a+b)^3}{ab}\right\}^{\frac{1}{3}}, \tag{2.3.3}$$

for large values of a and b. For a $Beta(a, b)$ prior, the larger a and b are, the more is the information contained in the prior. Hence, one should not choose too large values for a and b unless there is a lot of prior information. Thus, as one move towards a more informative prior from a vague prior, the ESSs obtained from the different methods yield similar results. To see how the ESS values differ for vague priors, the a and b parts of the ESS for small to moderate values of a and b were computed along with their asymptotic values (a and b part of (2.3.3)) and the prior standard deviations. The computations were repeated twice; once by considering $a = b$ and once by keeping them unequal. The results are respectively shown in Tables 2.2 and 2.3, and it can be seen from both that the difference between an ESS value and its asymptotic version is lower than 1.5. This clearly indicates that the difference between the methods are not too significant.

$a = b$	Prior SD	ESS_{VR} & ESS_{ELIR}	ESS_{PR}	ESS_{MTM}	$ESS_{MTM.P}$	ESS_{asym}
1.5	0.250	3.63	2.00	1.59	1.59	2.29
2.00	0.224	3.30	2.37	2.00	2.00	2.52
3.00	0.189	3.35	2.82	2.52	2.52	2.88
5.00	0.151	3.70	3.39	3.17	3.17	3.42
25	0.070	5.93	5.85	5.77	5.77	5.85
50	0.050	7.42	7.37	7.32	7.30	7.37
75	0.041	8.47	8.43	8.40	8.40	8.43

Table 2.2: Table showing the prior SD and a, b parts of different ESS (assuming $a = b$) along with their asymptotic value (column ESS_{asym})

a	b	Prior SD	ESS_{VR} & ESS_{ELIR}	ESS_{PR}	ESS_{MTM}	$ESS_{MTM.P}$	ESS_{asym}
1.5	2.5	0.217	3.63	2.25	1.95	2.20	2.57
2	3	0.200	3.42	2.59	2.28	2.38	2.75
4	8	0.131	4.10	3.69	3.50	3.62	3.78
20	35	0.064	6.29	6.18	6.11	6.13	6.19
50	40	0.052	7.20	7.14	7.09	7.09	7.14

Table 2.3: Table showing the prior SD and a, b parts of different ESS ($a \neq b$) along with their asymptotic value (column ESS_{asym})

2.4 Mixed Response Model with Independent Beta and Multivariate T Prior

In this section, we consider the same model as in (1.6.1), but instead of the bivariate normal prior for $\boldsymbol{\mu}$, we consider a Multivariate-T prior with d degrees of freedom for $\boldsymbol{\mu}$. The

advantage of using a Multivariate-T distribution over a Bivariate Normal distribution is that it is more robust to conflict between the data and the prior choice for the expectation of μ (Fan & Berger, 1992). Hence, we have

$$\begin{aligned} &p \sim Beta(a,b), (a>0, b>0) \text{ and} \\ &\boldsymbol{\mu} \sim T_2(d, \boldsymbol{\gamma}, \Sigma), \text{ independently of each other.} \end{aligned} \tag{2.4.1}$$

Here, $\boldsymbol{\gamma}$ and Σ are same as in (2.2.1). The likelihood and the information matrix will remain the same as in Section 2.2, but the posterior distribution of $\boldsymbol{\theta}$ will change. The posterior density of $\boldsymbol{\theta}$ is given by

$$\begin{aligned} p(\boldsymbol{\theta}|(y_i,z_i), i=1,2,\ldots,n) &\propto p(\boldsymbol{\theta}) L(\theta|(y_i,z_i), i=1,2,\ldots,n) \\ &\propto p^{a+n_1-1}(1-p)^{b+n_0-1}. \\ &\times \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(z_i - \mathbf{x}_i'\boldsymbol{\mu})^2\right\} . \frac{1}{\left(1+\frac{(\boldsymbol{\mu}-\boldsymbol{\gamma})'\Sigma^{-1}(\boldsymbol{\mu}-\boldsymbol{\gamma})}{d}\right)^{\frac{d+2}{2}}}. \end{aligned} \tag{2.4.2}$$

Thus, we see that the posterior distribution of p and $\boldsymbol{\mu}$ are independent, the posterior distribution of p is a $B(a+n_1, b+n_0)$, but the distribution of $\boldsymbol{\mu}$ cannot be expressed in the form of a standard distribution. The ESS for the different methods are calculated using the same formulae as in Table 2.1, and the results are reported in Table 2.4. The values

ESS	**Value**
ESS_{VR}	$\left\{\left(\frac{d-2}{d}\right)^2 \frac{\sigma^4}{\sigma_0^2\sigma_1^2(1-\rho^2)} . \frac{(a+b)(a+b-1)^2}{(a-1)(b-1)}\right\}^{\frac{1}{3}}$
ESS_{PR}	$\left\{\left(\frac{d-2}{d}\right)^2 \frac{\sigma^4}{\sigma_0^2\sigma_1^2(1-\rho^2)} . \frac{(a+b)^4(a+b+1)(a-1)(b-1)}{a^2b^2(a+b-1)(a+b-2)}\right\}^{\frac{1}{3}}$
ESS_{MTM}	$\left\{\left(\frac{d+2}{d}\right)^2 \frac{\sigma^4}{\sigma_0^2\sigma_1^2(1-\rho^2)} . (a+b)^2\left(\frac{a-1}{a^2}+\frac{b-1}{b^2}\right)\right\}^{\frac{1}{3}}$
$ESS_{MTM.P}$	$\left\{\left(\frac{d+2}{d}\right)^2 \frac{\sigma^4}{\sigma_0^2\sigma_1^2(1-\rho^2)} . \frac{(a+b-2)^3}{(a-1)(b-1)}\right\}^{\frac{1}{3}}$
ESS_{ELIR}	$\left\{\left(\frac{d}{d+4}\right)^2 \frac{\sigma^4}{\sigma_0^2\sigma_1^2(1-\rho^2)} . \frac{(a+b)(a+b-1)^2}{(a-1)(b-1)}\right\}^{\frac{1}{3}}$

Table 2.4: Formulas for calculating Multivariate ESS along with their values for model (1.6.1) and prior (2.4.1).

obtained are similar to the ones obtained in Table 2.1 with an extra factor added for the degrees of freedom. The same discussion as in Section 2.3, regarding the Beta prior goes for here as well, i.e. one should not choose too large values of a and b unless there is a lot of prior information. For the Multivariate-T part of the prior, we see that for fixed Σ, the variance of the prior decreases as d increases, and the prior becomes more informative. Since ESS is a method of quantifying information on the prior, its value should get higher as more and more information is incorporated into the prior. Therefore, as d increases

the ESS value should increase, which is true for ELIR, PR and VR methods, but not for the rest (see Table 2.7 for more details). ESS_{MTM} and $ESS_{MTM.P}$ will be much higher than ESS_{VR}, ESS_{PR} and ESS_{ELIR} if d is small. However, all the methods will provide similar values if a, b and d are large, i.e, when the prior is more informative.

2.5 Mixed Response Model with Unknown Variance for the Distribution of $Z|Y$

We consider the model as in (1.6.1), but σ^2 is assumed to be unknown in this case.

Here $\boldsymbol{\theta} = (p, \mu_0, \mu_1, \sigma^2)'$ is the collection of parameters of interest and we take suitable priors on $\boldsymbol{\theta}$ for finding the ESS.

We assume independent priors on p and $(\boldsymbol{\mu}, \sigma^2)$. Therefore, the prior of θ is given by

$$\begin{aligned} p &\sim Beta(a, b), a > 0, b > 0 \\ \boldsymbol{\mu}|\sigma^2 &\sim N_2(\boldsymbol{\gamma}, \sigma^2\Sigma), \text{ and} \\ \sigma^2 &\sim IG\left(\frac{c}{2}, \frac{d}{2}\right). \end{aligned} \tag{2.5.1}$$

Here, $\boldsymbol{\gamma}$ and Σ are same as in (2.2.1) and p and $(\mu_0, \mu_1, \sigma^2)'$ are independent.

The likelihood is given by,

$$L(\boldsymbol{\theta}|(y_i, z_i), i = 1, 2, \ldots, n) \propto p^{n_1}(1-p)^{n_0} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(z_i - \boldsymbol{x}'\boldsymbol{\mu})^2\right\}.$$

The information matrix based on a single observation is

$$i_F(\boldsymbol{\theta}) = \begin{bmatrix} \frac{1}{p(1-p)} & 0 & 0 & 0 \\ 0 & \frac{1-p}{\sigma^2} & 0 & 0 \\ 0 & 0 & \frac{p}{\sigma^2} & 0 \\ 0 & 0 & 0 & \frac{1}{2\sigma^4} \end{bmatrix},$$

and the posterior distribution of $\boldsymbol{\theta}$ comes out to be

$$\begin{aligned} p|(y_i, z_i), i = 1, 2, \ldots, n &\sim Beta(a^*, b^*), \\ \boldsymbol{\mu}|\sigma^2, (y_i, z_i), i = 1, 2, \ldots, n &\sim N_2(\boldsymbol{\gamma}^*, \sigma^2\Sigma^*), \text{ and} \\ \sigma^2|(y_i, z_i), i = 1, 2, \ldots, n &\sim IG\left(\frac{c^*}{2}, \frac{d^*}{2}\right), \end{aligned} \tag{2.5.2}$$

where $\boldsymbol{\gamma}^* = \begin{pmatrix}\gamma_0^*\\ \gamma_1^*\end{pmatrix}$, $\Sigma^* = \begin{bmatrix}\sigma_0^{*^2} & \rho^*\sigma_0^*\sigma_1^*\\ \rho^*\sigma_0^*\sigma_1^* & \sigma_1^{*^2}\end{bmatrix}$, $a^* = a + n_1$, $b^* = b + n_0$ and

$$\gamma_0^* = \frac{\left\{\left(n_1 + \frac{1}{\sigma_1^2(1-\rho^2)}\right)\left(\sum_{i=1}^n z_i(1-y_i) + \frac{\gamma_0}{\sigma_0^2(1-\rho^2)} - \frac{\rho\gamma_1}{\sigma_0\sigma_1(1-\rho^2)}\right) + \frac{\rho}{\sigma_0\sigma_1(1-\rho^2)}\left(\sum_{i=1}^n z_i y_i + \frac{\gamma_1}{\sigma_1^2(1-\rho^2)} - \frac{\rho\gamma_0}{\sigma_0\sigma_1(1-\rho^2)}\right)\right\}}{\left(n_0 + \frac{1}{\sigma_0^2(1-\rho^2)}\right)\left(n_1 + \frac{1}{\sigma_1^2(1-\rho^2)}\right) - \frac{\rho^2}{\sigma_0^2\sigma_1^2(1-\rho^2)^2}},$$

$$\gamma_1^* = \frac{\left\{\left(n_0 + \frac{1}{\sigma_0^2(1-\rho^2)}\right)\left(\sum_{i=1}^n z_i y_i + \frac{\gamma_1}{\sigma_1^2(1-\rho^2)} - \frac{\rho\gamma_0}{\sigma_0\sigma_1(1-\rho^2)}\right) + \frac{\rho}{\sigma_0\sigma_1(1-\rho^2)}\left(\sum_{i=1}^n z_i(1-y_i) + \frac{\gamma_0}{\sigma_0^2(1-\rho^2)} - \frac{\rho\gamma_1}{\sigma_0\sigma_1(1-\rho^2)}\right)\right\}}{\left(n_0 + \frac{1}{\sigma_0^2(1-\rho^2)}\right)\left(n_1 + \frac{1}{\sigma_1^2(1-\rho^2)}\right) - \frac{\rho^2}{\sigma_0^2\sigma_1^2(1-\rho^2)^2}},$$

$$\sigma_0^{*^2} = \frac{n_1 + \frac{1}{\sigma_1^2(1-\rho^2)}}{\left(n_0 + \frac{1}{\sigma_0^2(1-\rho^2)}\right)\left(n_1 + \frac{1}{\sigma_1^2(1-\rho^2)}\right) - \frac{\rho^2}{\sigma_0^2\sigma_1^2(1-\rho^2)^2}},$$

$$\sigma_1^{*^2} = \frac{n_0 + \frac{1}{\sigma_0^2(1-\rho^2)}}{\left(n_0 + \frac{1}{\sigma_0^2(1-\rho^2)}\right)\left(n_1 + \frac{1}{\sigma_1^2(1-\rho^2)}\right) - \frac{\rho^2}{\sigma_0^2\sigma_1^2(1-\rho^2)^2}},$$

$$\rho^* = \frac{\rho}{\sqrt{(1 + n_0\sigma_0^2(1-\rho^2))(1 + n_1\sigma_1^2(1-\rho^2))}},$$

$$c^* = c + n,$$

$$d^* = d + \sum_{i=1}^n (z_i - (1-y_i)\bar{z}_2 - y_i\bar{z}_1)^2 + \frac{1}{1-\rho^2}\left\{\frac{n_0}{\sigma_0^2}\left(n_1 + \frac{1}{\sigma_1^2}\right)(\bar{z}_2 - \gamma_0)^2 + \frac{n_1}{\sigma_1^2}\left(n_0 + \frac{1}{\sigma_0^2}\right)(\bar{z}_1 - \gamma_1)^2 - 2n_1n_0\frac{\rho}{\sigma_0\sigma_1}(\bar{z}_2 - \gamma_0)(\bar{z}_1 - \gamma_1)\right\}.$$

Note that the posterior distribution of p and $(\mu_0, \mu_1, \sigma^2)'$ are independent and the prior is conjugate. Its derivation has been omitted as it is similar to that shown in the Appendix A.1.

Analogous to (2.3.1), the ESS in this case is defined by,

$$ESS = n\{|Var(\theta)|^{-1}|E(Var(\hat{\theta}_n))|\}^{\frac{1}{4}}, \tag{2.5.3}$$

where $\hat{\theta}_n$ is an estimator of θ based on a sample of size n.

Again, we consider $\hat{\theta}_n$ to be the MLE of $\boldsymbol{\theta}$, i.e. $\hat{\theta}_n = (\hat{p}_n, \hat{\mu}_{0n}, \hat{\mu}_{1n}, \hat{\sigma}_n^2)'$, where

$$\begin{aligned}
\hat{p}_n &= \frac{1}{n}\sum_{i=1}^{n} y_i, \\
\hat{\mu}_{0n} &= \frac{\sum_{i=1}^{n}(1-y_i)z_i}{\sum_{i=1}^{n}(1-y_i)}, \\
\hat{\mu}_{1n} &= \frac{\sum_{i=1}^{n} y_i z_i}{\sum_{i=1}^{n} y_i}, \\
\hat{\sigma}_n^2 &= \frac{1}{n}\sum_{i=1}^{n}(z_i - (1-y_i)\hat{\mu_0}_n - y_i\hat{\mu_1}_n)^2.
\end{aligned}$$

Using (2.5.3), the ESS comes out to be approximately equal to

$$ESS \approx \left\{(c-2)\frac{1}{\sigma_0^2\sigma_1^2(1-\rho^2)}.\frac{(a+b)(a+b-1)^2}{(a-1)(b-1)}\right\}^{\frac{1}{4}}, \text{ provided } a>1, b>1, c>2. \tag{2.5.4}$$

Table 2.5 shows the ESSs calculated for prior (2.5.1) using different methods.

ESS	**Value**
ESS_{VR}	$\left\{(c-2).\frac{1}{\sigma_0^2\sigma_1^2(1-\rho^2)}.\frac{(a+b)(a+b-1)^2}{(a-1)(b-1)}\right\}^{\frac{1}{4}}$
ESS_{PR}	$\left\{\frac{(c-2)^4(c-4)}{c^3(c+2)}.\frac{1}{\sigma_0^2\sigma_1^2(1-\rho^2)}.\frac{(a+b)^4(a+b+1)(a-1)(b-1)}{a^2b^2(a+b-1)(a+b-2)}\right\}^{\frac{1}{4}}$
ESS_{MTM}	$\left\{(c-8).\frac{1}{\sigma_0^2\sigma_1^2(1-\rho^2)}.(a+b)^2\left(\frac{a-1}{a^2}+\frac{b-1}{b^2}\right).\frac{1}{1+2.\frac{b}{a+b}.\sigma_0^2+2.\frac{a}{a+b}.\sigma_1^2}\right\}^{\frac{1}{4}}$
$ESS_{MTM.P}$	$\left\{c.\frac{1}{\sigma_0^2\sigma_1^2(1-\rho^2)}.\frac{(a+b-2)^3}{(a-1)(b-1)}\right\}^{\frac{1}{4}}$
ESS_{ELIR}	$\left\{c.\frac{1}{\sigma_0^2\sigma_1^2(1-\rho^2)}.\frac{(a+b)(a+b-1)^2}{(a-1)(b-1)}\right\}^{\frac{1}{4}}$

Table 2.5: Multivariate ESS values for model (1.6.1) and prior (2.5.1).

Here ESS_{MTM} yields values different from the rest of the methods. ESS_{VR} and ESS_{ELIR} are approximately equal to the actual ESS given by (2.5.4). The same argument as in Section 2.3 and 2.4, regarding the hyper-parameters a and b remains valid in this case. For the inverse gamma prior we can say that as c increases, the prior variance of σ^2 decreases, provided the rest of the hyper-parameters are kept constant. Hence, c should not be chosen too big unless there is a lot of prior information. A similar table as in Table 2.2 comparing the c parts of the ESS was computed here, which showed that the values do not significantly differ even for small values of c and the differences became negligible as c increased. All the methods (except MTM) yield similar values which is approximately equal to $\left\{\frac{1}{\sigma_0^2\sigma_1^2(1-\rho^2)}.\frac{c(a+b)^3}{ab}\right\}^{\frac{1}{4}}$ for an informative prior with large values of a, b and c. Overall, we can say that in general when the prior becomes more informative,

the ESS obtained from the different methods are higher and the discrepancy between the different methods reduce, which should be the case.

2.6 Simulation Study

As it is already mentioned in Section 2.1.1, an ESS should preferably be predictively consistent. It is obvious from the formulas used for calculating ESS that the predictive consistency criteria will not be met in this case. The best we can do is to compare the expected posterior ESS subtracted by n with the prior ESS for different values of n and observe its deviation from the predictive consistency criteria.

2.6.1 Mixed Response Model with Independent Beta and Bivariate Normal Prior

The posterior ESS obtained using Model (1.6.1) and prior (2.2.1) is nothing but the prior ESS obtained by taking a $Beta(a+n_1, b+n_0)$ prior for p and $N_2(\gamma_0^*, \gamma_1^*, {\sigma_0^*}^2, {\sigma_1^*}^2, \rho^*)$ $[(\gamma_0^*, \gamma_1^*, {\sigma_0^*}^2, {\sigma_1^*}^2, \rho^*)'$ as in (2.2.2)$]$ for $(\mu_0, \mu_1)'$. The expressions for the expected values of the posterior ESS are in mathematically intractable form, and hence, we resort to simulation to obtain them. For the simulation, we take $p = 0.6, \sigma^2 = 64, \mu_0 = 0, \mu_1 = 1.5$ and generate the data $\{(y_i, z_i), i = 1, 2, \ldots, n\}$ $10,000$ times. For every generated set of $\{(y_i, z_i), i = 1, 2, \ldots, n\}$, we compute the posterior ESSs and take their average and standard deviation to get an estimate of the expected posterior ESS and its standard error. Next, we subtract the estimated expected posterior ESS by n and compare it with the prior ESS.

Table 2.6 shows the Prior ESS and Average Posterior $ESS(-n)$ for different values of n. The standard errors are reported within parentheses. The choice of the hyper-parameters $a, b, \gamma_0, \gamma_1, \sigma_0^2, \sigma_1^2, \rho$ are respectively 100, 75, -1, 1, 4, 9, 0.9. Other choices of

		Average Posterior ESS - n			
Method	Prior ESS	n= 10	n= 100	n= 1000	n= 10000
VR	75.66	89.16(0.020)	121.48(0.019)	143.71(0.011)	148.09(0.004)
PR	75.34	88.77(0.020)	120.88(0.019)	142.96(0.010)	147.32(0.004)
MTM	75.06	88.41(0.019)	120.35(0.019)	142.31(0.010)	146.65(0.004)
MTM.P	75.08	88.44(0.020)	120.41(0.019)	142.41(0.011)	146.76(0.004)
ELIR	75.66	89.16(0.020)	121.48(0.019)	143.71(0.011)	148.09(0.004)

Table 2.6: Table showing the Prior ESS and Average Posterior ESS-n(Standard Error) obtained for Model (1.6.1) and prior (2.2.1) taking $a = 100, b = 75, \gamma_0 = -1, \gamma_1 = 1, \sigma_0^2 = 4, \sigma_1^2 = 9, \rho = 0.9$.

p, σ^2, μ_0, μ_1 were also attempted but the results are not reported for the sake of brevity.

2.6.2 Mixed Response Model with Independent Beta and Multivariate T Prior

The prior (2.4.1) is not conjugate. However, if d is large, prior (2.4.1) can be well approximated by the conjugate prior (2.2.1) and their ESS-values are similar. This can be easily seen from Table 2.7 which reports the ESS-values obtained for different values of d in prior (2.4.1) keeping the other parameter-values the same as in Table 2.6. We see that as the degrees of freedom (d) increases, the ESS values approach the ESS values obtained using prior (2.2.1), which should be the case.

However, for small values of d, the prior cannot be approximated by a conjugate prior and the calculation of the Posterior ESS for a given dataset is not straightforward. Also, the posterior distribution is not in standard form. Therefore, to obtain the posterior mean, mode, variance etc, we need to generate observations from the posterior distribution. Samples are generated from the posterior distribution using the Metropolis-Hastings algorithm. In a similar manner as before, we generate the data $\{(y_i, z_i), i = 1, 2, \ldots, n\}$ $10,000$ times taking $p = 0.6, \sigma^2 = 64, \mu_0 = 0, \mu_1 = 1.5$, and compute the Posterior ESS for every generated dataset. The average Posterior ESS $(-n)$ along with its standard error is obtained and compared with the Prior ESS.

Table 2.8 shows the Prior ESS and Average Posterior ESS $-n$ for different values of n. The choice of the hyper-parameters $a, b, d, \gamma_0, \gamma_1, \sigma_0^2, \sigma_1^2, \rho$ are respectively, $100, 75, 5, -1, 1, 4, 9, 0.9$. We see that Prior ESS_{MTM} and $ESS_{MTM.P}$ are significantly higher than others, but the difference reduces when we consider Average Posterior $ESS - n$, especially when n is larger.

2.6.3 Mixed Response Model with the Variance of Z Conditioned on Y Unknown

Like the prior (2.2.1), the prior (2.5.1) is also conjugate, and hence, for a fixed dataset, the ESS calculated using the posterior density (2.5.2) is the ESS obtained by taking a $Beta(a^*, b^*)$ prior for p, a $N_2(\gamma_0^*, \gamma_1^*, \sigma_0^{*^2}, \sigma_1^{*^2}, \rho^*)$ for $(\mu_0, \mu_1)'$ and a $IG(\frac{c^*}{2}, \frac{d^*}{2})$ prior for σ^2 (where $a^*, b^*, \gamma_0^*, \gamma_1^*, \sigma_0^{*^2}, \sigma_1^{*^2}, \rho^*$ are as in (2.5.1)).

The Average Posterior ESS $(-n)$ along with its standard error are calculated in a similar manner as before, and compared with the Prior ESS. The model parameters for generating the data are taken as, $p = 0.6, \sigma^2 = 64, \mu_0 = 0, \mu_1 = 1.5$ and the hyper-parameters of the prior distribution are taken as $a = 100, b = 75, c = 100, d = 200, \gamma_0 =$

d	VR	PR	MTM	MTM.P	ELIR
3	36.37	36.22	105.51	105.54	43.01
5	53.82	53.60	93.93	93.96	51.13
10	65.20	64.93	84.76	84.78	60.46
20	70.53	70.23	79.98	80.01	67.00
50	73.63	73.32	77.04	77.07	71.87
100	74.65	74.33	76.05	76.08	73.71
200	75.15	74.84	75.56	75.58	74.67
500	75.46	75.14	75.26	75.28	75.26
1000	75.56	75.24	75.16	75.18	75.46
Bivariate Normal	75.66	75.34	75.06	75.08	75.66

Table 2.7: Table showing the ESS values obtained for different values of d using Model (1.6.1) and prior (2.4.1). The last row shows the ESS values obtained using Model (1.6.1) and prior (2.2.1). The values of the other hyper-parameters are taken as $a = 100, b = 75, \gamma_0 = -1, \gamma_1 = 1, \sigma_0^2 = 4, \sigma_1^2 = 9, \rho = 0.9$.

		Average Posterior ESS – n			
Method	Prior ESS	n=10	n= 100	n= 1000	n= 10000
VR	53.82	76.01(0.061)	113.98(0.162)	-	-
PR	53.6	75.68(0.061)	113.46(0.162)	-	-
MTM	93.93	92.25(0.126)	116.80(0.158)	126.50(0.205)	124.12(0.111)
MTM.P	93.96	92.28(0.126)	116.80(0.158)	126.56(0.205)	124.23(0.111)
ELIR	51.13	68.03(0.067)	102.58(0.092)	123.23(0.168)	125.13(0.108)

Table 2.8: Table showing the Prior ESS and Average Posterior ESS-n (Standard Error) obtained for Model (1.6.1) and prior (2.4.1) taking $a = 100, b = 75, d = 5, \gamma_0 = -1, \gamma_1 = 1, \sigma_0^2 = 4, \sigma_1^2 = 9, \rho = 0.9$.

$-1, \gamma_1 = 1, \sigma_0^2 = 4, \sigma_1^2 = 9, \rho = 0.9$. The values of the Prior ESS and Average Posterior ESS $-n$ are shown in Table 2.9. As expected, Tables 2.6, 2.8 and 2.9 show that the prior and posterior ESS-values are, in general, low when less information is available through the prior. However, it may be difficult to frame a general rule in this context. Table 2.9 is a more realistic scenario since it assumes that the variance of $Z|Y = y$ is unknown. In this setup, the different versions of ESS (except MTM) are more in agreement. But, we see that the performance of the methods were equally poor when it came to evaluating the deviation from the predictive consistency criteria. The 'Average Posterior ESS $-n$' moved further and further away from the 'Prior ESS' as the number of samples (n) increased, which shows that it contains information about the data, thereby violating the predictive consistency criterion.

		Average Posterior $ESS - n$			
Method	Prior_ESS	n= 10	n= 100	n= 1000	n= 10000
VR	10.09	30.24(0.013)	53.78(0.011)	66.37(0.006)	68.56(0.002)
PR	9.76	29.03(0.013)	51.17(0.011)	62.94(0.006)	64.99(0.002)
MTM	5.04	26.48(0.015)	50.54(0.011)	62.87(0.006)	64.98(0.002)
MTM.P	10.08	30.21(0.013)	53.61(0.011)	65.95(0.006)	68.07(0.002)
ELIR	10.14	30.43(0.013)	54.17(0.011)	66.86(0.006)	69.06(0.002)

Table 2.9: Table showing the Prior ESS and Average Posterior ESS-n(Standard Error) obtained for Model (1.6.1) (assuming unknown σ^2) and prior (2.5.1) taking $a = 100, b = 75, c = 100, d = 200, \gamma_0 = -1, \gamma_1 = 1, \sigma_0^2 = 4, \sigma_1^2 = 9, \rho = 0.9$.

2.7 Illustration with a Real-Life Vaccine Trial Dataset

In this section, we introduce a double-blind, randomized, multi-centre, phase II trial that has a binary safety and continuous efficacy data. The information from this data has been used as historical information, and the prior has been selected using this information. The prior selected may serve as external information for the future phase III trial. Ella et al. (2021) reports the interim findings of a double-blind, randomized, multi-centre, phase II trial on the safety and immunogenicity of an inactivated SARS-CoV-2 vaccine, BBV152. BBV152 is a whole-virion inactivated SARS-CoV-2 vaccine (3 μg or 6 μg) formulated with a toll-like receptor 7/8 agonist molecule (IMDG) adsorbed to alum (Algel). Participants were randomly assigned (1:1) to receive either 3 μg with Algel-IMDG or 6 μg with Algel-IMDG. A total of 190 patients were assigned to each group. Two intramuscular doses of vaccine were administered on day 0 and day 28. The primary outcome was SARS-CoV-2 wild-type neutralising antibody titres and seroconversion rates (defined as a post-vaccination titre that was at least four-fold higher than the baseline titre) at 4 weeks after the second dose (day 56), measured by use of the plaque-reduction neutralisation test (PRNT50) and the microneutralisation test (MNT50). The primary outcome was assessed in all participants who had received both doses of the vaccine. Safety was assessed by monitoring the adverse events in all participants who received at least one dose of the vaccine. Here $\log_{10}(titre)$ can be assumed to be normally distributed and hence we can consider $\log_{10}(titre)$ to be our measure of continuous efficacy and the occurrence of an adverse event as our measure of binary safety.

We only consider the data of the 3 μg with Algel-IMDG group for our illustration purpose. A similar analysis was also done using the data of the 6 μg with Algel-IMDG group but we do not report it here.

We define :

Y = 1 or 0 according as the patient in the 3 µg with Algel-IMDG group had any solicited local adverse reaction after dose 2, or not, and

$Z = \log_{10}(titre)$ of a patient in the 3 µg with Algel-IMDG group at day 56 measured using $PRNT_{50}$

Now we can apply model (1.6.1) to the data. However, since the individual participant data is not readily accessible, we have to rely on the summary measures reported in Ella *et al.* (2021) to set our prior. It has only been reported that 8 out of the total 189 participants in the 3 μg with Algel-IMDG group receiving dose 2 had solicited local adverse reactions. That means that $\sum_{i=1}^{n} y_i = 8$ where $n = 189$. Similarly, the Geometric Mean Titre ($GMT(PRNT_{50})$) at day 56, on the basis of the 184 patients analysed for immunogenicity has been reported as 100.9 ($95\%CI$ $74.1 - 137.4$). From this, we obtained the mean and standard deviation of z as $\bar{z} = 2, s_z = 0.93$. Note that, some of the z_i's are missing. This is due to the fact that safety was assessed for all the 189 participants receiving dose 2, but only 184 were analysed for immunogenicity as some were lost during follow-up. The mean and standard deviation have been obtained on the basis of the 184 available observations. For the joint modelling of the data as in (1.6.1), we need the (bivariate level) summary measure $\sum_{i=1}^{n} y_i z_i$, which is unavailable. This was one of the major challenges faced while handling the data. To overcome this challenge, we need to redesign the data by some sensible way. In particular, we resorted to simulation after making some assumptions. The first assumption we make is about the choice of the common variance σ^2. We assume that σ^2 is known, and set it to be equal to s_z^2. Since σ^2 is known, we assume an independent Beta and Bivariate Normal prior (prior (2.2.1)) for the data. The hyper-parameters $a, b, \gamma_0, \gamma_1, \sigma_0, \sigma_1, \rho$ are chosen using the information from the data. To select the values of a and b, we set the prior mean $(\frac{a}{a+b})$ to be equal to the observed proportion, $\hat{p} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{8}{189}$. Thus, we set $b = \left[\frac{1-p}{p}\right] + 1$, and consequently we have $a = b.\frac{p}{1-p}$. Forcing b to be greater than 1 guarantees that the prior mode would lie between 0 and 1. Having a prior mode at 1 would mean that we do not believe that $Y = 0$ is a very likely scenario. In this situation ($Y = 0$), there may not be any reason to conduct a clinical trial. Also, b should be large enough for a to be larger than 1. Otherwise, there is a mode at 0 and 1, and the ESS is undefined.

To set the values of $\gamma_0, \gamma_1, \sigma_0, \sigma_1, \rho$, simulated values of μ_0, μ_1 have been used. Since $\bar{z_1} = \frac{\sum_{i=1}^{n} y_i z_i}{\sum_{i=1}^{n} y_i}$ and $\bar{z_0} = \frac{\sum_{i=1}^{n} (1-y_i) z_i}{\sum_{i=1}^{n} (1-y_i)}$ are unavailable, it has been selected using simulation. Here n (= 189) independent and identically distributed (iid) observations have been generated from $N(\bar{z}, s_z^2)$ and a specific percentage (say, 25%) of the smallest values are selected. This specified percentage is called the cut-off value, and it is made to vary from 0.05 to 0.25, at an interval of 0.05. From the selected smallest values, an SRSWOR of size $\sum_{i=1}^{n} y_i (= 8)$ is selected, and its average is taken to be $\bar{z_1}$. Therefore, average of the rest of the values represent $\bar{z_0}$. The $\bar{z_0}$ and $\bar{z_1}$ obtained serve as simulated values of μ_0 and μ_1 respectively. In a similar manner, $10,000$ simulated values of μ_0 and μ_1 are generated for every cut-off value. To check if the algorithm makes sense, the range and inter-quartile range (IQR) of the generated μ_0 and μ_1 are computed for every cut-off value. Since the

range and IQR values do not vary much with the cut-off, we set the prior for (μ_0, μ_1) using the simulated values of μ_0 and μ_1. The values of $\gamma_0, \gamma_1, \sigma_0, \sigma_1, \rho$ are set by computing the mean vector and variance-covariance matrix of the generated $(\mu_0, \mu_1)'$. The prior ESSs obtained for each cut-off value are reported in Table 2.10.

	Prior ESS				
Cut-off	VR	PR	MTM	MTM.P	ELIR
0.05	595.74	40.49	58.62	561.83	595.74
0.1	606.61	41.23	59.69	572.09	606.61
0.15	611.56	41.56	60.17	576.76	611.56
0.2	578.57	39.32	56.93	545.64	578.57
0.25	598.90	40.70	58.93	564.82	598.90

Table 2.10: Table showing the Prior ESSs obtained using different methods for the different cut-off values.

From Table 2.10, we observe that for a given method, the cut-off values do not affect the values of the prior ESSs much. But, the ESSs obtained using the different methods vary widely. This is due to the fact that the value of the hyper-parameter a is small. Although a is small, the value of b is 23, and the variance of the Beta prior is 0.0016, which is quite small. Therefore, there is some information in the prior, and the value of ESS should be high. Hence we see that, MTM.P, VR and ELIR methods work best as they have high values and VR and ELIR provide values equal to the actual ESS given by (2.3.2), as mentioned earlier. The prior chosen here is dependent on the assumptions made. Of course, the algorithm could be improved, and the prior could be made more reliable if the entire dataset were available.

2.8 Concluding Remarks

In this chapter, we tried to calculate the Prior ESSs for the datasets where one of the components of the treatment-response vector is binary and the other is continuous. These kinds of mixed response datasets are often encountered in Phase II clinical trials. In such a situation, it is crucial to quantify the prior information in terms of Prior ESS. In general, we have seen that as we move towards a more informative prior, the values of the ESSs obtained from the different methods become more or less similar. But, as seen in Sections 2.2 to 2.5, the ESSs may vary widely when the prior is vague. VR and ELIR provided more reliable values in such cases.

Sometimes, due to lack of availability of prior information, one has to use a non-informative prior. The next question may arise, "How do we choose a non-informative prior?" One way to obtain vague prior from a normal distribution is to let the variance tend to infinity. For Beta distributions, let its parameters tend to 1. For example, if

we let $\sigma_0^2 = \sigma_1^2 = \kappa$ and $a = b = 1 + \frac{1}{\kappa}$ in prior (2.2.1), then for large κ the prior becomes non-informative and the ESS goes to $\left\{\frac{\sigma^4}{(1-\rho^2)}\right\}^{\frac{1}{3}}$ as $\kappa \longrightarrow \infty$. In Table 2.10, we observed that the ESSs varied widely because the value of one of the hyper-parameters was small. In such scenarios, VR and ELIR methods provided better results. It is always recommended to compute and compare every ESS as we do not know which one is going to perform better in a particular situation. If the calculation of every ESS is computationally challenging, it is advisable to go for the VR and ELIR methods.

Even though we have compared the Prior ESS and the Average Posterior ESS $(-n)$, the current methods do not satisfy the predictive consistency criterion. Development of methods that satisfy the predictive consistency criterion in case of mixed response models that provide reliable values in case of non-informative priors require further research. Nevertheless, the ESSs obtained in this chapter serves our purpose and we concentrate on the actual problem of sample size determination in the succeeding chapters.

Chapter 3

Sample Size Determination in Vaccine Trials

3.1 Introduction

The primary objective in the phase III clinical trial of any vaccine is to determine whether a vaccine is effective in preventing a particular disease, and if so, to what extent. The extent measured by means of protective vaccine efficacy, defined by VE (equation (1.2.1)). VE is usually estimated by using both cohort and case-control designs.

The randomized controlled trials conducted for a new vaccine application are usually comparative cohort designs where the data is obtained after monitoring the vaccinated and unvaccinated individuals for several months. From the data obtained in a vaccine trial, VE is estimated by presuming (1.3.1) to be the underlying model. On the other hand, case-control studies estimate VE with the help of OR and assume model (1.4.1) in the analysis. Nevertheless, both the models can be used in their respective designs to get the expressions for the standard error of $\widehat{VE}$ and CIs of VE.

The problem of sample size determination in vaccine trials was partially addressed in O'Neill (1988), where the required sample sizes were calculated when the width of the CI of VE was specified. However, another aspect of the sample size determination problem, which deals with the appropriate (optimal) selection of the allocation proportion ρ, was left unattended in the paper. It was assumed that an equal number of individuals were allocated across the two groups ('vaccine' and 'placebo') in comparative cohort designs. For case-control designs, the number of controls was taken to be c times the number of cases. The value of c was fixed and taken to be 1 or 4. The required sample size came out to be a function of the desired width of the CI, the probability of coverage, the assumed VE, and for cohort designs, the disease attack rate in the general population (ARU), and

for case-control designs, the assumed prevalence of vaccine exposure within the controls.

It has already been mentioned in Chapter 1 that several cost, logistic, and other ethical considerations in cohort designs have motivated researchers to deviate from the conventional equal allocation techniques towards suitable unequal designs. In this chapter, we try to calculate the required sample sizes when there is an uneven allocation of subjects across the groups instead of allocating an equal number of individuals to each group. The uneven allocation ratios derived here were compared to the standard equal allocation technique and the double allocation technique mentioned in Section 1.3. The concept was then extended to case-control studies where the quantity c was no longer assumed to be fixed.

Rest of the chapter is organized as follows. In Section 3.2, we calculate the sample sizes required in cohort studies when there is an unequal allocation of individuals across the vaccine and placebo groups. In Section 3.3, we will show that the proportion allocated to placebo ρ can be so chosen that some relevant criteria (like the total sample size, or the expected number of individuals developing the disease) is optimized, without compromising the precision of the study. In Section 3.4, we calculate the optimal sample sizes by fixing power of a relevant level α test instead of precision. In Section 3.5, we try to calculate the optimal sample sizes required for case-control studies. In Section 3.6, we deal with the data from a real-life trial and calculate the sample sizes using information from the trial. Section 3.7 shows how to implement our methods in a trial. Section 3.8 concludes.

3.2 Sample Sizes in Comparative Cohort Trials with Unequal Allocation

Here, we consider the binomial model described in Section 1.3 where the placebo and vaccine group sizes are n_A and n_B respectively. Using the delta method, one can obtain the asymptotic distribution of $\widehat{\psi}$ and $\widehat{\beta} = \log \widehat{\psi}$ as

$$\begin{aligned} \widehat{\psi} &\overset{a}{\sim} N\left(\psi, \frac{p_B^2(1-p_A)}{p_A^3} + \frac{p_B(1-p_B)}{p_A^2}\right), \\ \text{and } \widehat{\beta} &\overset{a}{\sim} N\left(\beta, \frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}\right), \end{aligned} \tag{3.2.1}$$

where $\sigma_A^2 = \frac{1-p_A}{p_A}$, $\sigma_B^2 = \frac{1-p_B}{p_B}$ (Koopman, 1984; Katz et al., 1978).

Consequently, the asymptotic $100(1-\alpha)\%$ CI of β is given by

$$\widehat{\beta} \mp z_{\alpha/2}\widehat{\sigma}.$$

Hence, O'Neill (1988) derived the $100(1-\alpha)\%$ CI for VE as

$$\left[1-\exp(\widehat{\beta}+z_{\alpha/2}\widehat{\sigma}), 1-\exp(\widehat{\beta}-z_{\alpha/2}\widehat{\sigma})\right].$$

Here,

$$\sigma = \sqrt{\frac{\sigma_A^2}{n_A}+\frac{\sigma_B^2}{n_B}}, \tag{3.2.2}$$

and $\widehat{\sigma}$ is its estimate obtained by replacing p_A and p_B in σ by their respective estimates $\widehat{p}_A$ and $\widehat{p}_B$.

Subsequently, the width $W(\widehat{\beta},\widehat{d})$ and the relative width (with respect to VE) $RW(\widehat{\beta},\widehat{d})$ of the CI for VE were obtained as

$$W(\widehat{\beta},\widehat{d}) = (1-\widehat{VE})(e^{\widehat{d}}-e^{-\widehat{d}}), \tag{3.2.3}$$

$$RW(\widehat{\beta},\widehat{d}) = \frac{(1-\widehat{VE})}{\widehat{VE}}(e^{\widehat{d}}-e^{-\widehat{d}}), \tag{3.2.4}$$

where $\widehat{d} = z_{\alpha/2}.\widehat{\sigma}$. The asymptotic width/relative width was fixed to some value, and the minimum sample size required to achieve that width was taken as the required sample size.

Under the assumption of equal numbers N of vaccinated and unvaccinated subjects (i.e. $n_A = n_B = N$), the value of N was obtained as

$$N = \left(\frac{z_{\frac{\alpha}{2}}}{d}\right)^2\left(\frac{1-p_A}{p_A}+\frac{1-p_B}{p_B}\right),$$

where d is the desired half-width of the CI on the log scale, obtained from the specified W/RW by solving equation (3.2.3) or (3.2.4).

Thus, the total sample size required is $n = 2N$, which is a function of the specified width/relative width, the vaccine efficacy, and the attack rate in the unvaccinated population.

As mentioned earlier, keeping the cost and other constraints in clinical trials in mind, it is sometimes advisable to include more individuals to the vaccine group. Thus, instead of allocating N individuals to each group, we try to calculate the sample sizes required when there are an uneven allocation of subjects across the groups. We know $\frac{n_A}{n} = \rho$, i.e. a fraction ρ of the incoming individuals are allocated to placebo, and the rest are allocated to the vaccine group.

If d is the desired half-width of the CI for β, then by fixing $d(= z_{\frac{\alpha}{2}}\sigma)$ and assuming $n_A = n\rho$, we get the required sample size as

$$n = \left(\frac{z_{\frac{\alpha}{2}}}{d}\right)^2 \left\{\frac{1-p_A}{p_A\rho} + \frac{1-p_B}{p_B(1-\rho)}\right\}. \tag{3.2.5}$$

One can obtain another form of the CI of VE, viz.

$$(1-\widehat{\psi}) \mp z_{\frac{\alpha}{2}} \left\{\frac{\widehat{p}_B(1-\widehat{p}_A)}{\widehat{p}_A^3} + \frac{\widehat{p}_B(1-\widehat{p}_B)}{\widehat{p}_A^2}\right\},$$

by directly using the asymptotic distribution of $\widehat{\psi}$ given by (3.2.1). If one fixes the asymptotic length of this CI to some quantity l, then the minimum sample size required to achieve that width is given by

$$n = 4\left(\frac{z_{\frac{\alpha}{2}}}{l}\right)^2 \cdot \left\{\frac{p_B(1-p_A)}{p_A^3\rho} + \frac{p_B(1-p_B)}{p_A^2(1-\rho)}\right\}. \tag{3.2.6}$$

Formulas (3.2.5) and (3.2.6) can be used interchangeably, and they eventually produce nearly similar results. Formula (3.2.5) has been developed along the same line as O'Neill (1988) and will be used in the remainder of the chapter. The importance of formula (3.2.6) will be clearer in Chapter 4 while comparing it with the Bayesian approach. However, it should be noted that the optimal ρ's derived in the succeeding section (Section 3.3) will remain unchanged regardless of the formula used.

Overall it can be said that if the allocation proportion ρ is given, one can easily calculate the required sample size, provided the anticipated vaccine efficacy and the disease attack rate in the unvaccinated population are known a priori.

3.3 Choice of ρ

The primary purpose of diverting from the equal allocation setup is to overcome the cost and other ethical constraints mentioned in Chapter 1. This is achieved by choosing the allocation proportion ρ suitably (Brittain & Schlesselman, 1982; Ivanova & Rosenberger, 2001). In other words, ρ must be so chosen that within the given level of precision, some statistically relevant criteria are optimal. In this chapter, we consider the total sample size and the expected number of people developing the disease as two possible criteria. If some other criterion is of interest, then the corresponding optimal allocation proportion ρ can be derived in a similar manner.

3.3.1 Neyman Allocation

Here our objective is to choose ρ in such a way that the total number of individuals recruited is minimized keeping the precision intact. As in Section 3.2, we fix the desired level of precision by specifying d, or equivalently σ^2 (since $d = z_{\alpha/2}\sigma$). Thus, if we fix σ^2 to k, a prefixed constant, then the optimum ρ will minimize the total sample size

$$\begin{aligned} &n = n_A + n_B \text{ such that} \\ &\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B} = k. \end{aligned} \tag{3.3.1}$$

This yields the optimum choice of ρ as

$$\begin{aligned} \rho(p_A, p_B) &= \frac{\sigma_A}{\sigma_A + \sigma_B} = \frac{\sqrt{\frac{q_A}{p_A}}}{\sqrt{\frac{q_A}{p_A}} + \sqrt{\frac{q_B}{p_B}}} \\ &= \frac{\sqrt{q_A p_B}}{\sqrt{q_A p_B} + \sqrt{q_B p_A}}, \end{aligned} \tag{3.3.2}$$

where $q_A = 1 - p_A$ and $q_B = 1 - p_B$.

If, instead of $\widehat{\beta}$, the variance of $\widehat{p}_B - \widehat{p}_A$ was fixed to k, we would have obtained Neyman's allocation (Melfi & Page, 1998; Ivanova & Rosenberger, 2001; Biswas & Bhattacharya, 2016). Due to this analogy, we call it 'Neyman allocation' in our discussion.

3.3.2 RSIHR Allocation

Investigators are often interested in an allocation that reduces, on an average, the total number of individuals developing the disease within the study. This gives rise to our second criterion which is the expected number of individuals developing the disease. As before, we fix σ^2 to k and the optimum ρ in this case will minimize

$$\begin{aligned} &E = n_A p_A + n_B p_B \text{ such that} \\ &\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B} = k. \end{aligned} \tag{3.3.3}$$

Here, the optimum choice of ρ comes out to be

$$\rho(p_A, p_B) = \frac{\sqrt{q_A} p_B}{\sqrt{q_A} p_B + \sqrt{q_B} p_A}. \tag{3.3.4}$$

This has similarity to the RSIHR allocation (Rosenberger et al., 2001; Zhang & Rosenberger, 2006), which minimizes the expected number of treatment failures while testing for treatment difference. We simply call it RSIHR allocation.

If the investigator specifies the vaccine efficacy and the disease attack rate in the unvaccinated population, one can obtain the optimum allocation proportion ρ under both Neyman and RSIHR allocation. If the the desired width/relative width of the CI and the probability of coverage $(1-\alpha)$ are also specified, then one can easily calculate the required total sample size (from equation (3.2.5)) under both Neyman and RSIHR allocations.

Tables 3.1 and 3.2 show the total sample size required to construct 95% CI for VE, when VEs are assumed to be 40% and 80%, respectively. The sample size of the corresponding vaccine group are also reported within parentheses. The relative widths are fixed at $1, 0.8, 0.6, 0.4$, and the disease attack rate in the unvaccinated population are taken as $0.1, 0.01, 0.001$. The sample sizes required under equal allocation scheme (when equal number of individuals are allocated to vaccine and placebo) O'Neill (1988) and double allocation scheme (when the number of individuals allocated to vaccine is twice that of placebo) are also reported along with the Neyman and RSIHR allocation schemes for comparison.

From Tables 3.1 and 3.2, it can be seen that Neyman allocation provides the lowest sample sizes, which is quite obvious from the way it has been derived. RSIHR provides slightly lesser sample sizes than equal allocation, but it allocates higher number of individuals to the vaccine group for higher vaccine efficacy. It ensures that, on an average, a lesser number of individuals develop the disease. The number of samples required increase drastically as the disease attack rate in the general population decreases. Also, a much higher number of samples are required if we use equal allocation instead of Neyman or RSIHR allocation. Double allocation provides the highest sample sizes when VE is 0.4, but the sample sizes are nearly equal to Neyman when $VE = 0.8$. This is due to the fact that the optimum Neyman's allocation ratio is close to the double allocation ratio in the latter case. But this may not always hold true.

As pointed out by O'Neill (1988) and Carpenter (2001), the sample size required to achieve a CI of a pre-specified width can change dramatically as a function of the disease attack rate in the unvaccinated population and the magnitude of the index VE. This remains true in case of Neyman and RSIHR allocation also, but the difference between the sample sizes required under the Neyman allocation and the equal allocation becomes wider as the VE increases. This can be seen from Table 3.3 which shows the total sample size and the proportion allocated to vaccine under different allocation schemes for different values of VE, when the width of the interval W is specified as 0.24 and ARU is specified as 0.01. The corresponding relative width and the confidence intervals are also shown.

Note that, the proportion of individuals allocated to vaccine under Equal and Double allocation schemes have not been reported, as they are fixed quantities taking values 0.5 and 0.67, respectively. It is clear that, as VE increases, the sample size decreases, and a much lesser sample size is required in case of Neyman allocation. For better comparison, the percentage reduction in sample size as compared to the Equal allocation scheme has been shown for the three other schemes. In case of both Neyman and RSIHR allocations, the percentage reduction in sample size is higher for a more efficacious vaccine. However, the magnitude of reduction is much lesser in case of RSIHR, as it focuses on the reduction of number of cases rather than sample size. Like Equal allocation, Double allocation is just another fixed allocation scheme, and the sample size as compared to equal allocation may or may not decrease, depending on the situation. In a nutshell, we can say that it is always advisable to use Neyman allocation over equal allocation, particularly when the disease attack rate in the general population is small and the vaccine is highly efficacious.

In clinical trials, the investigator may often be interested to know the minimum expected time (in weeks or days) that would be required to get an estimate of the vaccine efficacy at a desired level of precision. As seen in case of COVID trials, it is often of interest to minimize this time duration. Now suppose that people are recruited in the study at a constant rate, and the number of people arriving every week is a constant l. One may want to determine the optimum ρ that minimizes the number of weeks required to achieve a given level of precision. If l people are constantly recruited every week, the the number of weeks required to recruit $n(= n_A + n_B)$ people is

$$W = \left[\frac{n_A + n_B}{l}\right] + 1. \tag{3.3.5}$$

Therefore, minimizing W is equivalent to minimizing the total sample size $n = n_A + n_B$. Thus, the allocation ratio ρ that minimizes the expected time to achieve a given level of precision is nothing but the Neyman's allocation ratio. Hence, the Neyman allocation not only minimizes the total sample, but it also reduces the expected time required to get an answer from the study. The number of weeks that will be required for a study implementing RSIHR or Equal or Double can also be obtained using the formula (3.3.5). The number of weeks required for the study under the different allocation schemes are shown within parentheses, beside the corresponding sample sizes. It is assumed that people are coming in at a constant rate of 1000 per week. Evidently, Neyman allocation takes the smallest amount of time. RSIHR takes similar time as Equal allocation, but the additional time will be compensated by the fact that it reduces the overall disease incidence. Double Allocation takes the highest time when the vaccine efficacy is low, but the time reduces as the vaccine efficacy increases, and becomes nearly equal to the time required in Neyman allocation when the vaccine is highly efficacious. Therefore, to remain on the safer side, it is always advisable to go for Neyman allocation if our central goal is

to reduce time.

Again, it can be seen from Table 3.3 that higher the vaccine efficacy higher is the proportion allocated to the vaccine group for RSIHR. This should be the case as RSIHR tries to minimize the expected number of individuals developing the disease. So, if the vaccine is really effective in preventing the disease, RSIHR will try to allocate more and more individuals to the vaccine group, thereby satisfying our ethical considerations.

		ARU		
RW	Allocation Type	0.1	0.01	0.001
1	Equal	1767 (884)	18964 (9482)	190932 (95466)
	Double	1809 (1206)	19543 (13029)	196885 (131257)
	Neyman	1735 (987)	18658 (10523)	187896 (105890)
	RSIHR	1762 (1110)	18959 (11859)	190927 (119338)
0.8	Equal	2727 (1364)	29263 (14632)	294624 (147312)
	Double	2792 (1861)	30157 (20105)	303810 (202540)
	Neyman	2677 (1523)	28792 (16239)	289939 (163397)
	RSIHR	2720 (1714)	29256 (18299)	294616 (184149)
0.6	Equal	4800 (2400)	51508 (25754)	518581 (259290)
	Double	4914 (3276)	53081 (35387)	534750 (356500)
	Neyman	4711 (2680)	50677 (28582)	510335 (287603)
	RSIHR	4787 (3016)	51495 (32209)	518568 (324129)
0.4	Equal	10723 (5361)	115055 (57527)	1158373 (579187)
	Double	10977 (7318)	118569 (79046)	1194491 (796327)
	Neyman	10523 (5986)	113200 (63845)	1139954 (642430)
	RSIHR	10693 (6737)	115026 (71945)	1158344 (724020)

Table 3.1: Total sample size required to obtain 95% CI of specified relative width *RW* under different allocation schemes and different assumed attack rate in the unvaccinated population ARU. The vaccine efficacy VE is assumed to be at 40%. The corresponding sample sizes for the vaccine group are reported within parentheses.
Table generated using formulae (3.2.5)

3.3.3 Modified Neyman and RSIHR Allocation under Restricted Vaccine Dosage

At times, there may be a limited vaccine supply and the number of people that can be allocated to vaccine will be restricted. This happens when the cost of vaccination is high, or the resources are limited. For example, in an outbreak of Ebola disease caused by Sudan virus, only a few hundred of the vaccine were available, and implementing unequal allocation techniques (especially for RSIHR allocation) in such a case is challenging. However, such challenges can be overcome by slightly modifying the optimization problems

RW	Allocation Type	ARU 0.1	ARU 0.01	ARU 0.001
1	Equal	214 (107)	2205 (1102)	22111 (11056)
	Double	185 (123)	1927 (1285)	19346 (12897)
	Neyman	184 (129)	1922 (1329)	19294 (13333)
	RSIHR	211 (177)	2202 (1836)	22109 (18425)
0.8	Equal	286 (143)	2945 (1473)	29541 (14770)
	Double	247 (165)	2575 (1716)	25846 (17230)
	Neyman	246 (172)	2567 (1776)	25776 (17813)
	RSIHR	282 (236)	2941 (2453)	29537 (24616)
0.6	Equal	432 (216)	4451 (2226)	44645 (22322)
	Double	374 (249)	3891 (2594)	39060 (26040)
	Neyman	372 (261)	3880 (2684)	38956 (26921)
	RSIHR	426 (357)	4445 (3707)	44639 (37201)
0.4	Equal	830 (415)	8559 (4279)	85846 (42923)
	Double	719 (479)	7482 (4988)	75108 (50072)
	Neyman	716 (501)	7461 (5161)	74907 (51766)
	RSIHR	818 (687)	8547 (7128)	85834 (71533)

Table 3.2: Total sample size required to obtain 95% CI of specified relative width *RW* under different allocation schemes and different assumed attack rate in the unvaccinated population ARU. The vaccine efficacy VE is assumed to be at 80%. The corresponding sample sizes for the vaccine group are reported within parentheses.
Table generated using formulae (3.2.5)

(3.3.1) and (3.3.3). If only k' people can be allocated to vaccine, i.e. $n_B \leq k'$, then the addition of this constraint to equations (3.3.1) and (3.3.3), and solving the resulting the constrained optimization will give us our desired allocation proportion.

Under the limited dosage set-up, we minimize

$$\begin{aligned} & n = n_A + n_B \text{ subject to} \\ & \frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B} = k, \\ \text{and } & n_B \leq k' \end{aligned} \tag{3.3.6}$$

or, we minimize

$$\begin{aligned} & E = n_A p_A + n_B p_B \text{ subject to} \\ & \frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B} = k, \\ \text{and } & n_B \leq k' \end{aligned} \tag{3.3.7}$$

			Equal Allocation	Double Allocation		Neyman Allocation			RSIHR Allocation		
VE	RW	CI	Total Sample Size (Number of weeks)	Total Sample Size (Number of weeks)	Percentage reduction in sample size*	Total Sample Size (Number of weeks)	Proportion Allocated to vaccine	Percentage Reduction in sample size*	Total Sample Size (Number of weeks)	Proportion Allocated to vaccine	Percentage reduction in sample size*
0.8	0.3	(0.65,0.89)	14199 (15)	12413 (13)	12.58	12377 (13)	0.69	12.83	14180 (15)	0.83	0.13
0.6	0.4	(0.46,0.70)	30583 (31)	29463 (30)	3.66	29090 (30)	0.61	4.88	30563 (31)	0.71	0.07
0.4	0.6	(0.27,0.51)	51508 (52)	53081 (54)	-3.05	50677 (51)	0.56	1.61	51495 (52)	0.63	0.03
0.3	0.8	(0.17,0.41)	63581 (64)	67286 (68)	-5.83	63074 (64)	0.54	0.8	63573 (64)	0.59	0.01

Table 3.3: Table showing the total sample size and the proportion allocated to vaccine under the different allocation schemes for different values of VE when the width of the 95% CI is fixed at 0.24. The number of weeks required for the study under the assumption that people are coming in at a constant rate of 1000 per week are also shown within parentheses.

*Percentage reductions in sample sizes reported in comparison to Equal Allocation.

according to our objective.

The modified Neyman and RSIHR allocation obtained by solving (3.3.6) and (3.3.7), gives the best optimal ratio under the restricted dosage. It can be seen from Table 3.1 that when we set $RW = 1$, $ARU = 0.01$, and $VE = 0.4$, then the sample sizes under Equal, Double, Neyman, and RSIHR allocation are respectively calculated as 18964, 19543, 18658, and 18959. Now, if a maximum of $10,000$ doses of the vaccine are available ($k' = 10,000$), then the required sample size under the modified Neyman setup turns out to be 18726, which is still less than those obtained under Equal and Double allocation by an amount of 238 and 817 respectively. Given the fact that the cost per enrollment can range between 1000 and 5000 USD, a reduction in sample size of around $200 - 800$ can save thousands of dollars. A similar thing applies for modified RSIHR allocation as well. It ensures that even with a limited vaccine supply, more people are saved from the disease occurrence.

3.4 Sample Sizes with Fixed Power

In the previous section, we had fixed the precision at a given level and obtained the optimal sample sizes. There is another way of obtaining optimal sample size based on power. In this approach, we fix the power of the relevant test at a specified level, say 95%, and find the optimal allocation ratio that minimizes our relevant criteria. Based on the objective of the trial, one decides upon the approach to be followed. However, the two approaches are expected to behave similarly for large values of n (Brittain & Schlesselman, 1982).

If we want to test for the effectiveness of the vaccine in preventing the disease, the relevant hypothesis is given by,

$$H_0 : VE \leq 0 \text{ vs. } H_1 : VE > 0,$$

which can equivalently be written as,

$$H_0 : \beta \geq 0 \text{ vs. } H_1 : \beta < 0.$$

The Z-test based on large sample normal approximation of the ditribution of $\widehat{\beta}$ rejects H_0 at level α if $Z < -z_\alpha$, where

$$Z = \frac{\widehat{\beta}}{\widehat{\sigma}_0}$$

is the test statistic and $\widehat{\sigma}_0$ is the estimate of σ (as in (3.2.2)) under H_0. Note that,

$$\widehat{\sigma}_0 = \frac{1-\widehat{p}}{\widehat{p}}\left(\frac{1}{n_A}+\frac{1}{n_B}\right), \tag{3.4.1}$$

where $\widehat{p} = \frac{\widehat{p}_A n_A + \widehat{p}_B n_B}{n_A+n_B}$ is the common estimate of p_A and p_B under H_0.

Asymptotic formulas for the power and sample size can be derived using an approach similar to Chan & Bohidar (1998). The power function of the test is given by,

$$\begin{aligned} p(\beta) &= P_\beta\left(\frac{\widehat{\beta}}{\widehat{\sigma}_0} < -z_\alpha\right) \\ &= P_\beta\left(\frac{\widehat{\beta}-\beta}{\sigma} < \frac{-z_\alpha\widehat{\sigma}_0-\beta}{\sigma}\right) \\ &= P_\beta\left(Z < -\frac{z_\alpha\widehat{\sigma}_0+\beta}{\sigma}\right), \text{ where } Z \sim N(0,1). \end{aligned}$$

Therefore, the asymptotic power of the test is given by,

$$1-\Phi\left(\frac{z_\alpha\bar{\sigma}_0+\beta}{\sigma}\right), \tag{3.4.2}$$

where Φ is the Cumulative Distribution Function (CDF) of a $N(0,1)$ distribution and $\bar{\sigma}_0$ is the limiting value of $\widehat{\sigma}_0$ under the alternative β, obtained by replacing $\widehat{p}_A$ and $\widehat{p}_A$ in (3.4.1) by p_A and p_B.

If a fixed proportion ρ of the incoming individuals are allocated to placebo, then the expression for the asymptotic power in (3.4.2) can alternatively be expressed as,

$$1-\Phi\left(\frac{z_\alpha\sqrt{\frac{\rho q_A+(1-\rho)q_B}{\rho p_A+(1-\rho)p_B}}\sqrt{\frac{1}{n\rho(1-\rho)}}+\beta}{\sqrt{\frac{1}{n}}\sqrt{\frac{q_A}{p_A\rho}+\frac{q_B}{p_B(1-\rho)}}}\right)$$

If the power is fixed to $(1-\eta)$, then the sample size required to achieve that power (asymptotic expression) is given by,

$$\boxed{n = \left(\frac{z_\alpha\sqrt{\frac{\rho q_A+(1-\rho)q_B}{\rho p_A+(1-\rho)p_B}}\sqrt{\frac{1}{\rho(1-\rho)}}+z_\eta\sqrt{\frac{q_A}{p_A\rho}+\frac{q_B}{p_B(1-\rho)}}}{\beta}\right)^2}. \tag{3.4.3}$$

The sample size comes out to be a function of the anticipated values of VE and p_A (i.e., the values under the assumed alternative), the level of the test α, the desired level of

power $1-\eta$, and the allocation proportion ρ. Now, as in Section 3.3, we can find the optimum ρ that minimizes the total sample size $n_A + n_B$ or the expected number of people developing the disease $n_A p_A + n_B p_B$. To facilitate the comparison, we again call the optimum ρ obtained in the former case as Neyman allocation proportion, and that in the latter case as RSIHR allocation proportion. The optimum ρ in both the cases cannot be written in explicit form as in Section 3.3, but has been obtained numerically. However, it will be a function of the assumed values of p_A and VE. Therefore, similar to Section 3.3, the required sample sizes under both Neyman and RSIHR allocation can be easily obtained, provided the size of the Z-test, the desired level of power, the assumed values of VE and p_A are given.

Table 3.4 shows the sample sizes required to achieve 95% power under different allocation schemes for various values of p_A(orARU) when the underlying VEs are assumed to be 0.5 and 0.9. Note that the Neyman and RSIHR allocation are not the same as in (3.3.2) and (3.3.4), but have been obtained numerically. The optimal allocation proportions have also been reported within parenthesis. The equal and double allocation with $\rho = 0.5$ and $\rho = 1/3$ have also been reported for comparison. The conclusions drawn are similar to Section 3.3; Neyman allocation gives the smallest sample size and RSIHR allocation reduces the overall chance of a person developing the disease. But, the advantage of this approach is that it can be compared with the conditional binomial model given in Chan & Bohidar (1998).

3.4.1 Comparison with the Conditional Binomial Model

The conditional binomial model is often used in vaccine trials for sample size determination and inference. For large sample size and small disease attack rate, the number of cases of the disease may be assumed to follow Poisson distributions, in both vaccinated and unvaccinated populations, and an exact conditional binomial test can be devised (Chan & Bohidar, 1998; Gail, 1974). The test can be further used to get an idea about the expected number of samples to be used for the study. The expected sample size that will be required to achieve a desired power, say 95% in a level α conditional binomial test, can be obtained, and compared with the sample size required to achieve the same 95% power in the level α unconditional Z-test, described in the previous section.

Suppose n_A and n_B are sufficiently large, and p_A and p_B are sufficiently small such that $n_A p_A \to \lambda_A$ and $n_B p_B \to \lambda_B$. Then, it may be assumed that

$$x_A \sim Poisson(\lambda_A) \text{ and } x_B \sim Poisson(\lambda_B),$$

independently of each other. Then,

$$x_B|x_A + x_B = T \sim Bin(T, \kappa),$$

where $\kappa = \frac{\lambda_B}{\lambda_A+\lambda_B} = \frac{n_B p_B}{n_A p_A+n_B p_B} = \frac{1-VE}{1+c-VE}$ and $c = \frac{n_A}{n_B} = \frac{\rho}{1-\rho}$.

Testing $H_0 : VE \leq 0$ vs. $H_1 : VE > 0$ is equivalent to testing

$$H_0 : \kappa \geq \kappa_0 \text{ vs.} H_1 : \kappa < \kappa_0, \tag{3.4.4}$$

where $\kappa_0 = \frac{1}{1+c}$.

We reject H_0 at level α if $x_B \leq x_c$ where x_c is such that $P(x_B \leq x_c|x_B \sim Bin(T, \kappa_0)) \leq \alpha$. The power function of the test is given by,

$$p(\kappa) = P(x_B \leq x_c|x_B \sim Bin(T, \kappa)).$$

Again, if we fix the power to $1-\eta$, we can find the T that achieves this power provided the allocation ratio c (or ρ), the assumed values of VE and p_A under the alternative are known. Since, the unconditional expected value of T is $n_A p_A + n_B p_B$, once we get T we can easily get the expected number of samples required for the study using the relation

$$n \approx \frac{T(1+c)}{(c+1-VE)p_A}. \tag{3.4.5}$$

But, obtaining an optimal allocation ratio in this case is difficult as the value of κ under the null hypothesis (κ_0) is a function of c. Every time a different c is chosen, the hypothesis changes, and the sample size so obtained will be incomparable. The sample sizes will give the same power but the underlying test will be different. Obtaining an optimal sample size in such a situation requires further research. However, if c is fixed to some constant, likes $c = 1$ or $\frac{1}{2}$, we can easily obtain the sample size required to achieve a desired power, provided the level of the test α and the value of κ (through VE and p_A) under the alternative are known.

These sample sizes can be easily compared to the sample sizes obtained using the unconditional Z-test in the previous section. The expected total sample size that will be required if we fix the power of the conditional test to 95% are reported in the last column of Table 3.4. The values of c are taken to be 1 (equal allocation) or $\frac{1}{2}$ (double allocation). The assumed values of VE are 0.5 and 0.9, and ARUs are taken to be $0.1, 0.05$ and 0.01.

We see that the Neyman allocation provides the smallest sample size. The Equal and Neyman allocation for the unconditional Z-test gives equal sample sizes when $VE = 0.5$. This is due to the fact that the equal allocation ratio is the Neyman allocation ratio, for

VE	ARU	Unconditional Z-test				Conditional Binomial test	
		Equal allocation	Double allocation	Neyman allocation	RSIHR allocation	Equal allocation	Double allocation
0.5	0.1	1185(0.5)	1333(0.67)	1185(0.5)	1220(0.59)	1347(0.5)	1500(0.67)
0.5	0.05	2460(0.5)	2768(0.67)	2460(0.5)	2533(0.59)	2693(0.5)	3000(0.67)
0.5	0.01	12663(0.5)	14246(0.67)	12663(0.5)	13036(0.59)	13467(0.5)	15000(0.67)
0.9	0.1	270(0.5)	285(0.67)	267(0.55)	342(0.76)	273(0.5)	257(0.67)
0.9	0.05	548(0.5)	580(0.67)	542(0.55)	696(0.76)	545(0.5)	514(0.67)
0.9	0.01	2776(0.5)	2936(0.67)	2745(0.55)	3525(0.76)	2727(0.5)	2571(0.67)

Table 3.4: Table showing the total sample size required to achieve 95% power in the unconditional Z-test and the conditional binomial Test under different allocation schemes for various assumed values of VE and ARU. The proportion allocated to vaccine in the different allocation schemes are also shown in parenthesis beside the corresponding sample sizes.
Table generated using formulas (3.4.3) and (3.4.5)

this specific case. The RSIHR and Double allocation proportion provides higher sample sizes. Note that the double allocation in the conditional test provides higher sample sizes than it's unconditional counterpart when VE is 0.5. The scenario is reversed when $VE = 0.9$, and hence no general conclusion can be made regarding the superiority of one test over another, in terms of reduction of sample size. Although Equal allocation in the conditional set-up provides lesser sample size for both the VE values, the observations are again case-specific, and no conclusion can be drawn. However, Neyman allocation is expected to reduce the sample size to the largest extent. Hence, it is always safer to go for Neyman allocation if reduction of sample size is of primary interest.

3.5 Sample Sizes in Case-Control Studies with Unequal Allocation

In case-control studies, the model (1.4.1) is considered and vaccine effiacy is estimated with the help of OR. Here, n_A controls and n_B cases of the disease are taken, and the asymptotic distribution of the estimate of the OR, $\widehat{\xi}$ and $\widehat{\phi} = \log \widehat{\xi}$ are given by,

$$\widehat{\xi} \overset{a}{\sim} N\left(\xi, \frac{p_B^2(1-p_A)}{p_A^3(1-p_B)^2} + \frac{p_B(1-p_A)^2}{p_A^2(1-p_B)^3}\right), \tag{3.5.1}$$
$$\text{and } \widehat{\phi} \overset{a}{\sim} N\left(\phi, \frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}\right),$$

where $\sigma_A^2 = \frac{1}{p_A q_A}$ and $\sigma_B^2 = \frac{1}{p_B q_B}$ (O'Neill, 1984, 1988). Consequently, a confidence interval for ϕ is given by

$$\left(\widehat{\phi} - z_{\alpha/2}\widehat{\sigma}, \widehat{\phi} + z_{\alpha/2}\widehat{\sigma}\right),$$

where $\sigma^2 = \frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}$, and a CI for $VE = 1 - e^{\phi}$ is given by

$$\left(1 - e^{\widehat{\phi}+\widehat{d}}, 1 - e^{\widehat{\phi}-\widehat{d}}\right),$$

where $\widehat{d} = z_{\alpha/2}\widehat{\sigma}$.

As before, the relative width of the confidence interval for VE (with respect to VE) can be written as

$$RW = \frac{1 - VE}{VE}\left(e^{d} - e^{-d}\right). \tag{3.5.2}$$

If the relative width RW is specified, we can find the desired value of d by solving equation (3.5.2).

Let c denotes the control-to-case ratio, i.e. $c = n_A/n_B$. O'Neill (1988) showed that the number of cases required will be

$$n_B = \left(\frac{z_{\frac{\alpha}{2}}}{d}\right)^2 \left\{\frac{1}{p_B(1-p_B)} + \frac{1}{cp_A(1-p_A)}\right\}, \tag{3.5.3}$$

and hence

$$\boxed{n = \left(\frac{z_{\frac{\alpha}{2}}}{d}\right)^2 \left\{\frac{1}{p_B(1-p_B)} + \frac{1}{cp_A(1-p_A)}\right\}(1+c)} \tag{3.5.4}$$

Hence, the number of cases and the total sample size required are functions of the desired half-width (d) of the CI for ϕ, the probability of coverage α, the rate of prevalence of vaccine exposure in the control group (p_A), and the anticipated vaccine efficacy (VE).

Alternatively, if one fixes the asymptotic length of the CI for VE based on the asymptotic distribution of $\widehat{\xi}$ (equation (3.5.1)), viz.

$$(1 - \widehat{\xi}) \mp z_{\frac{\alpha}{2}}\left(\frac{p_B^2(1-p_A)}{p_A^3(1-p_B)^2} + \frac{p_B(1-p_A)^2}{p_A^2(1-p_B)^3}\right)$$

to l, then the required sample size will be given by,

$$\boxed{n = 4\left(\frac{z_{\frac{\alpha}{2}}}{l}\right)^2 \cdot \left(\frac{p_B^2(1-p_A)}{cp_A^3(1-p_B)^2} + \frac{p_B(1-p_A)^2}{p_A^2(1-p_B)^3}\right)(1+c).} \tag{3.5.5}$$

Equations (3.5.4) and (3.5.5) are two alternative formulas and either of them can

be used to get an insight about the sample size. The former expression will be used throughout the rest of the chapter while the latter will be of more relevance in Chapter 4.

3.5.1 Neyman Allocation

O'Neill (1988) considered the control-to-case ratio c to be equal to 1 or 4. Instead of fixing c arbitrarily, here we try to choose the optimum c that minimizes the total sample size $n = n_A + n_B$ provided the variance σ^2 (or, equivalently half-width d) is fixed to some value k. As in Section 3.3.1, this gives us an allocation analogous to the Neyman allocation when the test is carried out using $\widehat{\phi}$. The optimum c is obtained as

$$c(p_A, p_B) = \frac{\sigma_A}{\sigma_B} = \frac{\sqrt{p_B q_B}}{\sqrt{p_A q_A}}, \tag{3.5.6}$$

where $q_A = 1 - p_A$ and $q_B = 1 - p_B$.

Thus, if the vaccine efficacy and the prevalence rate of vaccine exposure in controls are specified, the optimum c under Neyman allocation can be obtained from equation (3.5.6). Subsequently, the total sample size and the number of cases required can be calculated using equation (3.5.4) and (3.5.3) if the desired relative width of the CI and the probability of coverage $1 - \alpha$ are also specified.

Since the case-control design is a retrospective design, the cases and controls are checked for their vaccination status. Hence, our ethical consideration of minimizing the expected number of people developing the disease is irrelevant over here. Therefore, developing an allocation analogous to the RSIHR allocation (as in Section 3.3.2) is not of much relevance.

Table 3.5 and 3.6 shows the total sample size n required to obtain 95% CI of relative width 1 and 0.5, when vaccine efficacies are respectively assumed to be 0.4 and 0.8. The prevalence rates of vaccine exposure in the controls are taken in the range of 0.1 to 0.5. Apart from considering Neyman allocation where the value of c is obtained from equation (3.5.6), we also consider Equal ($c = 1$) and Quadruple ($c = 4$) allocations due to their historical importance as those values were used by O'Neill (1988). The commonly used $c = 2$ (Double allocation) has also been considered.

Again, we see that Neyman allocation provides much lower sample sizes than the equal or quadruple allocation when the prevalence of vaccine exposure in controls is low and the vaccine efficacy is high. Thus, it is desirable to use Neyman allocation when a high precision is desired for a highly efficacious vaccine with low prevalence rate of vaccine

VE = 0.4								
RW = 1								
	Equal		Double		Quadruple		Neyman	
p_A	Cases	Total	Cases	Total	Cases	Total	Cases	Total
0.1	1010	2019	810	2431	711	3555	1105	1996
0.2	540	1080	428	1283	372	1859	582	1072
0.3	391	782	305	916	263	1314	414	778
0.4	325	650	250	751	213	1064	337	649
0.5	296	592	225	674	189	943	301	592
RW = 0.5								
	Equal		Double		Quadruple		Neyman	
p_A	Cases	Total	Cases	Total	Cases	Total	Cases	Total
0.1	3933	7865	3157	9472	2770	13848	4304	7776
0.2	2103	4206	1667	5000	1449	7243	2267	4175
0.3	1522	3045	1190	3570	1024	5120	1613	3032
0.4	1265	2531	975	2924	829	4146	1314	2527
0.5	1154	2307	875	2624	735	3675	1172	2307

Table 3.5: Total sample size (cases + controls) and the number of cases of the disease required when the relative width (w.r.t. VE) is fixed at 1 or 0.5 and the vaccine efficacy is assumed to be 40%.
Table generated using formulae (3.5.4)

exposure in controls. Although the total sample size is reduced for Neyman allocation, the control-to-case ratio comes out to be less than one in almost all the cases, especially when the vaccine efficacy is low. This is a disadvantage of Neyman allocation as it is always desirable to take a higher number of controls than cases.

3.6 Real-Life Example

In this section, we try to calculate the optimal sample sizes required when the values of the underlying vaccine efficacy (VE), the assumed disease attack rate in the general population (ARU), and the desired width/relative width (W/RW) are taken from the pooled results of four randomized clinical trials.

Voysey et al. (2021) reports the pooled interim findings from four blinded, randomized, controlled trials done across the UK, Brazil, and South Africa. Participants aged 18 years and older were randomly assigned (1:1) to ChAdOx1 nCoV-19 vaccine or control (meningococcal group A, C, W, and Y conjugate vaccine or saline). Participants in the ChAdOx1 nCoV-19 group received two doses containing 5×10^{10} viral particles (standard

VE = 0.8								
RW = 1								
	Equal		Double		Quadruple		Neyman	
p_A	Cases	Total	Cases	Total	Cases	Total	Cases	Total
0.1	107	214	97	291	92	459	129	191
0.2	52	104	46	139	44	218	62	95
0.3	34	68	30	89	28	138	40	64
0.4	25	51	22	65	20	98	29	49
0.5	21	41	17	51	15	76	23	40
RW = 0.5								
	Equal		Double		Quadruple		Neyman	
p_A	Cases	Total	Cases	Total	Cases	Total	Cases	Total
0.1	287	575	260	780	246	1231	346	514
0.2	140	280	124	373	117	584	167	256
0.3	92	183	80	239	74	369	108	172
0.4	68	136	58	174	53	264	79	131
0.5	55	111	45	136	41	203	62	108

Table 3.6: Total sample size (cases + controls) and the number of cases of the disease required when the relative width (w.r.t. VE) is fixed at 1 or 0.5 and the vaccine efficacy is assumed to be 80%.
Table generated using formulae (3.5.4)

dose; SD/SD cohort); a subset in the UK trial received a half dose as their first dose (low dose) and a standard dose as their second dose (LD/SD cohort). The primary efficacy analysis included symptomatic COVID-19 in seronegative participants with a nucleic acid amplification test-positive swab more than 14 days after a second dose of vaccine. The primary objective was to evaluate the efficacy of ChAdOx1 nCoV-19 vaccine against NAAT-confirmed COVID-19. The primary outcome was virologically confirmed, symptomatic COVID-19, defined as a NAAT-positive swab combined with at least one qualifying symptom (fever $\geq 37.8°C$, cough, shortness of breath, or anosmia or ageusia). There were 131 cases of symptomatic COVID-19 (primary outcome) in the 11,636 LD/SD or SD/SD recipients who were eligible for inclusion in the primary efficacy analysis more than 14 days after second dose of vaccine. Out of them, there were 30 (0.5%) cases among the 5807 participants in the vaccine arm and 101 (1.7%) cases among 5829 participants in the control group, resulting in a vaccine efficacy of 70.4%. Among the participants receiving two standard doses, vaccine efficacy was 62.1% (95% CI 41.0-75.7; 27 [0.6%] of 4440 in the ChAdOx1 nCoV-19 group vs 71 [1.6%] of 4455 in the control group) and in participants who received a low dose followed by a standard dose, efficacy was 90.0% (67.4-97.0; three [0.2%] of 1367 vs 30 [2.2%] of 1374). A total of 18 participants (7 in vaccine

group and 11 in control group) had other non-primary symptomatic COVID-19 disease. Other non-primary symptomatic COVID-19 disease includes cases who have symptoms other than the five main symptoms that are required for inclusion in the primary analysis (e.g., a participant who has diarrhoea and malaise but no fever, cough, shortness of breath, anosmia, or ageusia). Therefore, the vaccine efficacy including all symptomatic (primary + non-primary) COVID-19 cases was 67.1% (37 of 5807 in vaccine group vs. 112 of 5829 in control group). A total of 69 participants (29 in vaccine group, 40 in control group) were asymptomatic or the symptoms were unknown. The vaccine efficacy considering any NAAT-positive swab, which includes all the symptomatic and asymptomatic cases, was obtained as 55.7% (68 of 5807 vs. 153 of 5829).

For all the situations mentioned above, 95% CI for VE can also be obtained as the data on the occurrence of the number of cases and the number of allocations to each of the vaccine and control group are available. Hence, the width/relative width of the CI can be obtained subsequently. Now, if we reverse the problem and consider the obtained relative width to be our desired relative width, the estimated vaccine efficacy and the proportion of cases in the control group to be the true values of VE and ARU respectively, we can obtain the sample sizes required under the different allocation schemes (using equation (3.2.5)) and compare them with actual sample size of the trial. Table 3.7 computes these sample sizes under the four allocation schemes (Equal, Double, Neyman, and RSIHR) using the estimates from the five different scenarios explained in the previous paragraph. The columns VE, ARU and RW show the estimated VE, ARU and RW of the 95% CI obtained from the trial, which serve as the assumed VE, ARU and desired RW while computing the sample sizes under the Equal, Double, Neyman and RSIHR allocation schemes. The total sample sizes required under different schemes are shown in the Equal, Double, Neyman and RSIHR columns. The actual sample size, i.e. the sample size taken during the trial, is also reported for comparison. The corresponding sample sizes for the vaccine arm are also shown within parentheses. Table 3.7 shows that if Neyman allocation is used instead of the actual (1:1) allocation, the same relative width would have been achieved by a much lesser number of samples, provided the estimated VE and ARU remain the same. RSIHR allocation provides slightly lesser sample sizes than the actual one, and also allocates more individuals to the vaccine group, thereby reducing the overall chance of a person developing the disease. We see that the double allocation provides nearly equal sample sizes as the Neyman allocation. This happens because the n_A and n_B obtained by solving the optimization problem (equation (3.3.1)) in this case is close to that of double allocation. For example, in the case of SD/SD recipients we have considered $p_A = 0.016$ and $p_B = 0.006$ in our calculations. This gives the optimal ρ from equation (3.3.2) as 0.38, which is quite close to $\frac{1}{3}$ – the fixed ρ in case of double allocation. Similar phenomena occurred in the other four situations (see Table 3.7) as well. Now, suppose p_A and p_B were respectively 0.01 and 0.006, then the optimal ρ would have been

0.44 (closer to $\frac{1}{2}$) and the sample sizes would have been closer to equal allocation instead.

3.7 Implementation

It should be noted that the numerical values of vaccine efficacy (VE) and the disease attack-rate in the general population (ARU) are unknown, and are to be estimated from the trial. Hence, at the beginning of the trial, we have no idea regarding the sample size and the desired allocation proportion (Neyman or RSIHR or anything else). If the individuals are being recruited sequentially, we do not know when to stop recruiting or how to allocate the incoming individuals to the vaccine or placebo group so that our desired allocation is achieved. Even though we do not have an estimate of the sample, we can still achieve Neyman or RSIHR allocation by resorting to sequential methods like SMLE, DBCD and ERADE procedures . If we assume that the participants are entering the trial in a sequential manner and the information on all the previous patients are available when a new patient is enrolled to the study, SMLE/DBCD/ERADE procedures may be well-applied to the data. In general, SMLE/DBCD/ERADE procedures target a specified allocation proportion ρ to one of the two treatments. In our case, the target allocation proportion will be the allocation ratio $\rho = \rho(p_A, p_B) = \frac{n_A}{n}$ to the placebo group. If we wish to allocate the individuals according to Neyman (or RSIHR) allocation, then our ρ will be as in equation (3.3.2) (or equation (3.3.4)), so that for large n, the actual allocation proportion will approach the desired allocation proportion.

Suppose participants are entering the trial one by one and they are assigned to vaccine or placebo. After a specified time, each individual is checked for the occurrence of the disease (response). Let $\widehat{p}_A(j-1)$ and $\widehat{p}_B(j-1)$ denote the estimate of p_A and p_B after $(j-1)$ individuals have responded. Then, under the SMLE procedure, the jth individual is assigned to placebo with probability $\rho(\widehat{p}_A(j-1), \widehat{p}_B(j-1))$.

Let $N_A(j-1)$ denotes the number of individuals assigned to placebo out of the first $(j-1)$ individuals. Then, under the DBCD procedure (Eisele, 1994; Eisele & Woodroofe, 1995), the jth individual is assigned to placebo with probability

$$g\left(\frac{N_A(j-1)}{j-1}, \rho(\widehat{p}_A(j-1), \widehat{p}_B(j-1))\right),$$

where $g : [0,1] \times [0,1] \longrightarrow [0,1]$ is a function satisfying some regularity conditions. The conditions impose certain continuity, differentiability and monotonicity restrictions on the function g as well as on ρ. Eisele (1994) and Eisele & Woodroofe (1995) suggested a specific form of g, and Melfi et al. (2001) pointed out that the form itself did not satisfy those restrictive conditions. Hu & Zhang (2004) later on relaxed a few of the assumptions, and the four following conditions were considered on g and ρ:

	VE	ARU	RW	Actual Sample Size	Equal Allocation	Double Allocation	Neyman Allocation	RSIHR Allocation
All LD[a]/SD[b] and SD/SD receipients	0.702	0.0173	0.3544	11636 (5807)	11624 (5812)	10701 (7134)	10685 (6926)	11605 (8952)
SD/SD recipients	0.618	0.0159	0.5623	8895 (4440)	8888 (4444)	8494 (5663)	8409 (5208)	8878 (6435)
LD/SD recipients	0.899	0.0218	0.3311	2741 (1367)	2735 (1368)	2235 (1490)	2149 (1636)	2724 (2478)
Any symptomatic COVID-19 disease	0.668	0.0192	0.3754	11636 (5807)	11625 (5812)	10869 (7246)	10823 (6884)	11606 (8730)
Any NAAT-positive swab	0.554	0.0262	0.4626	11636 (5807)	11627 (5814)	11383 (7589)	11169 (6716)	11611 (8047)

Table 3.7: Table showing the total sample size that would have been required under the different allocation schemes if we wanted to achieve the relative width obtained from the study. The estimated vaccine efficacy and disease attack-rate in the unvaccinated population from the study serve as the assumed value of VE and ARU while calculating sample sizes. The actual sample sizes used in the study are also shown for comparison, and the corresponding sample sizes for the vaccine arm are shown within parentheses.

[a]LD: Low Dose
[b]SD: Standard Dose

(i) g is jointly continuous

(ii) $g(x,x) = x$ for all $x \in [0,1]$

(iii) $g(x,y)$ is strictly decreasing in x and strictly increasing in y

(iv) ρ is a continuous function and it is twice continuously differentiable in a small neighborhood of (p_A, p_B).

We choose the function g of the form suggested by Hu & Zhang (2004), i.e.

$$g(x,\rho) = \begin{cases} 1 & \text{if } x = 0 \\ \frac{\rho\left(\frac{\rho}{x}\right)^{\alpha_1}}{\rho\left(\frac{\rho}{x}\right)^{\alpha_1} + (1-\rho)\left(\frac{(1-\rho)}{(1-x)}\right)^{\alpha_1}} & \text{if } 0 < x < 1 \\ 0 & \text{if } x = 1. \end{cases} \tag{3.7.1}$$

where α_1 is a positive integer controlling the randomness of the procedure. Note that, for $\alpha_1 = 0$, we get the SMLE procedure (Rosenberger et al., 2001; Zhang et al., 2006).

ERADE (Hu et al., 2009) is another RAR procedure. It allocates the jth individual to the placebo group with probability

$$p_j = \begin{cases} \alpha_2\widehat{\rho}_{j-1}, & \text{if } \frac{N_A(j-1)}{j-1} > \widehat{\rho}_{j-1} \\ \widehat{\rho}_{j-1} & \text{if } \frac{N_A(j-1)}{j-1} = \widehat{\rho}_{j-1} \\ 1 - \alpha_2(1-\widehat{\rho}_{j-1}) & \text{if } \frac{N_A(j-1)}{j-1} < \widehat{\rho}_{j-1}, \end{cases} \tag{3.7.2}$$

where $\widehat{\rho}_{j-1} = \rho(\widehat{p}_A(j-1), \widehat{p}_B(j-1))$ and $0 \leq \alpha_2 < 1$ is a constant that reflects the degree of randomization.

Hu & Rosenberger (2003) showed the precise link between the power and variability of the design. The lesser the variability, i.e. the lesser the value of $n.Var\left(\frac{N_A(n)}{n}\right)$, the higher the power and the better the design.

The variability of the DBCD and ERADE procedures with Neyman and RSIHR as target allocation proportions, obtained from the results given in Hu & Zhang (2004) and Ivanova (2003) can be easily obtained. Note that the variances will be explicit functions of p_A and p_B. The curve of the asymptotic variances $(n.Var\left(\frac{N_A(n)}{n}\right))$ against p_B (when p_A has been fixed to 0.01) of the various RAR procedures, with Neyman and RSIHR as target allocation proportion, has respectively been shown in Figures 3.1 and 3.2.

For further illustration, the asymptotic variances for the SMLE, DBCD ($\alpha_1 = 2$) and ERADE procedures with Neyman and RSIHR as respective target allocation proportions for different values of p_A and VE are shown in Tables 3.8 and 3.9. The disease attack rates in the unvaccinated population are taken in the range of 0.1 to 0.9 with a difference

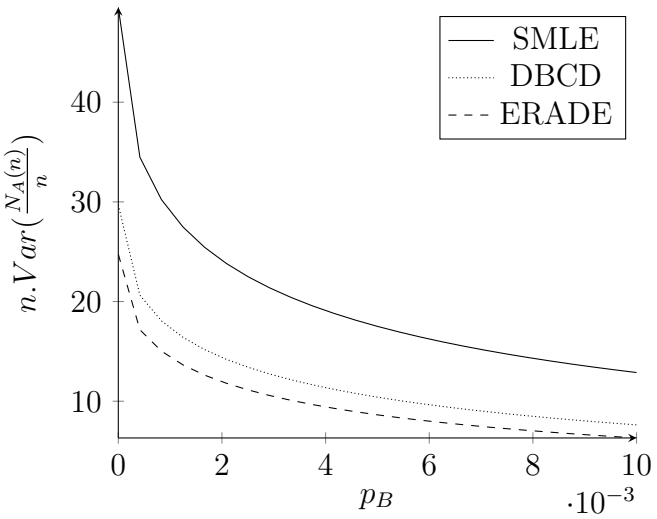

Figure 3.1: Plot of p_B vs. the asymptotic variances ($n.Var(\frac{N_A(n)}{n})$) of the SMLE , DBCD and ERADE procedure with Neyman as the target allocation proportion. The value of p_A has been fixed at 0.01.

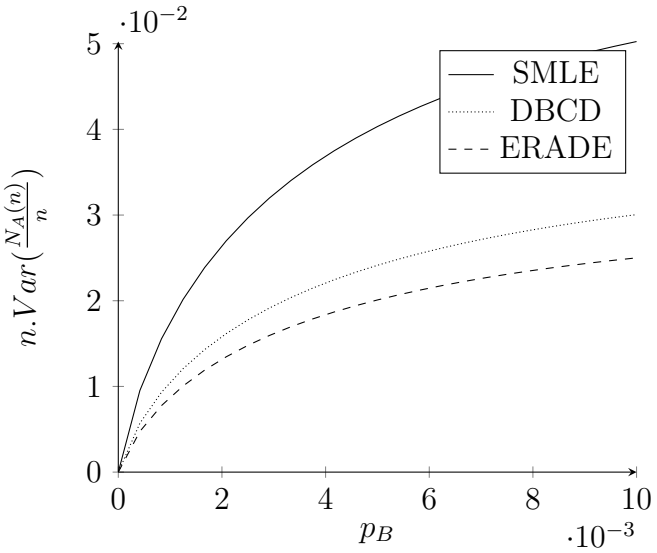

Figure 3.2: Plot of p_B vs. the asymptotic variances ($n.Var(\frac{N_A(n)}{n})$) of the SMLE , DBCD and ERADE procedure with RSIHR as the target allocation proportion. The value of p_A has been fixed at 0.01.

of 0.1, and the vaccine efficacies are in the range of 0.1 to 0.9 with 0.2 difference. It is already a well-known fact that the ERADE procedure provide best results among the three. It provides the least variances, followed by DBCD, and then SMLE procedure (Hu & Rosenberger, 2003; Rosenberger & Hu, 2004). Figures 3.1, 3.2 and Tables 3.8, 3.9 provides further evidence in support of this. Another RAR procedure that provides asymptotically best results is the Drop-the-Loser (DL) urn design (Ivanova, 2003). But it cannot be implemented as it targets a specified allocation ratio and cannot be used to target any desired allocation ratio. Although the ERADE procedure provides a better design than DBCD or SMLE, any one of the three can be applied in vaccine trials to

achieve a desired allocation proportion if the implementation of the better one is difficult.

		Incidence proportion in unvaccinated population (p_A)								
VE		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	SMLE	1.70	1.06	0.86	0.78	0.75	0.76	0.81	0.95	1.36
	DBCD	0.92	0.54	0.42	0.37	0.35	0.36	0.39	0.47	0.72
	ERADE	0.73	0.41	0.31	0.27	0.25	0.26	0.28	0.35	0.56
0.3	SMLE	1.86	1.13	0.90	0.80	0.75	0.74	0.77	0.86	1.17
	DBCD	1.02	0.58	0.44	0.38	0.35	0.35	0.37	0.42	0.62
	ERADE	0.81	0.44	0.33	0.28	0.25	0.25	0.27	0.31	0.48
0.5	SMLE	2.07	1.22	0.95	0.82	0.75	0.72	0.73	0.79	1.02
	DBCD	1.15	0.64	0.47	0.40	0.36	0.34	0.35	0.39	0.54
	ERADE	0.92	0.49	0.36	0.29	0.26	0.25	0.25	0.29	0.42
0.7	SMLE	2.39	1.35	1.01	0.84	0.74	0.69	0.67	0.69	0.85
	DBCD	1.34	0.72	0.52	0.42	0.36	0.33	0.33	0.35	0.45
	ERADE	1.08	0.56	0.39	0.31	0.27	0.24	0.24	0.26	0.35
0.9	SMLE	3.02	1.57	1.09	0.84	0.70	0.60	0.54	0.51	0.57
	DBCD	1.74	0.87	0.59	0.44	0.36	0.30	0.27	0.26	0.31
	ERADE	1.42	0.70	0.46	0.34	0.27	0.23	0.21	0.20	0.24

Table 3.8: Asymptotic variances of the SMLE, DBCD ($\alpha_1 = 2$) and ERADE procedure with Neyman allocation as target for different values of VE and ARU.

3.7.1 Drawbacks in Implementation

The sample sizes required for Neyman and RSIHR allocations (in Tables 3.1, 3.2 and 3.4) have been calculated by assuming that p_A and p_B (or, equivalently VE and ARU) values are known upfront, and do not account for the adaptive estimation of them. However, it is clear that p_A and p_B are unknown, and to achieve Neyman or any such allocation ratio in real life, we have to rely on RAR methods such as SMLE, DBCD or ERADE. However, there are some drawbacks that the RAR techniques suffer. In Equal and Double allocation schemes, the target allocation ratios are independent of p_A and p_B, and thus they need not be estimated adaptively. Therefore, if we conduct a real life trial and stop at the point when the desired precision is achieved, it is expected that the finally achieved sample sizes will be, on an average, similar to the ones calculated theoretically in Tables 3.1, 3.2 and 3.4. But, in case of RAR designs, the target allocation proportions vary with the values of p_A and p_B, and must be estimated adaptively. Therefore, in such a case, the sample sizes are necessarily random, and the average sample size might not match exactly with the calculated sample sizes. The statistical power and estimated precision are also random in RAR designs, and depend on the variability of the chosen RAR procedure.

		Incidence proportion in unvaccinated population (p_A)								
VE		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
	SMLE	5.51	2.90	2.05	1.63	1.41	1.28	1.24	1.30	1.65
0.1	DBCD	3.21	1.64	1.13	0.88	0.75	0.67	0.65	0.68	0.89
	ERADE	2.63	1.33	0.90	0.69	0.58	0.52	0.50	0.53	0.70
	SMLE	5.93	3.09	2.15	1.69	1.43	1.27	1.18	1.18	1.37
0.3	DBCD	3.46	1.76	1.19	0.92	0.76	0.67	0.62	0.62	0.74
	ERADE	2.84	1.42	0.96	0.73	0.60	0.52	0.48	0.48	0.59
	SMLE	6.08	3.12	2.13	1.64	1.35	1.16	1.04	0.99	1.07
0.5	DBCD	3.56	1.78	1.19	0.90	0.73	0.62	0.55	0.52	0.58
	ERADE	2.93	1.45	0.96	0.71	0.57	0.48	0.43	0.41	0.46
	SMLE	5.48	2.74	1.83	1.37	1.09	0.91	0.78	0.70	0.70
0.7	DBCD	3.22	1.58	1.03	0.76	0.59	0.49	0.41	0.37	0.39
	ERADE	2.65	1.29	0.83	0.60	0.47	0.38	0.32	0.29	0.31
	SMLE	2.92	1.41	0.91	0.65	0.50	0.40	0.32	0.27	0.26
0.9	DBCD	1.72	0.82	0.52	0.36	0.27	0.21	0.17	0.15	0.14
	ERADE	1.42	0.67	0.42	0.29	0.22	0.17	0.14	0.12	0.11

Table 3.9: Asymptotic variances of the SMLE, DBCD ($\alpha_1 = 2$) and ERADE procedure with RSIHR allocation as target for different values of VE and ARU.

These drawbacks remain true for any other allocation technique that depends on p_A and p_B, and has to be implemented using RAR methods. To verify how RAR affects the Neyman and RSIHR allocation ratio in our case, we performed several simulation studies.

In the simulation trials that were conducted, we allocated the incoming individuals to either vaccine or placebo group targeting one of Equal, Double, Neyman or RSIHR allocation ratios. To achieve equal allocation, simulation trials were conducted where every incoming individual was allocated to either vaccine or placebo with equal probability ($1:1$ randomization). Similarly, double allocation was achieved by allocating incoming individuals to the vaccine and placebo groups with respective probabilities $2/3$ and $1/3$ ($2:1$ randomization). For each of Neyman or RSIHR allocation, three simulations were conducted where individuals were allocated to vaccine or placebo using the three RAR schemes viz. SMLE, DBCD and ERADE, with the relevant allocation ratio as target. The responses were generated by assuming a true VE of 0.4 and ARU of 0.1. The desired width of the confidence interval for VE was taken to be 0.5.

For every combination of allocation ratio and RAR/non-RAR procedure, 1000 identical trials were conducted that targeted the said allocation ratio, and allocated individuals using the relevant RAR/non-RAR randomization procedure. In each trial, the estimate of VE and its confidence interval were obtained at every stage, till the width of the CI becomes less than 0.5. The total number of individuals and the number allocated to

vaccine till the attainment of the desired width, and thier averages give us an idea about the expected number of sample sizes that would be required when we actually conduct the trial. The numbers have been reported in the first part of Table 3.10. The sample sizes for the vaccine group have been reported within parentheses. These numbers were then compared to the sample sizes calculated using formula (3.2.5), reported in the first column of Table 3.10. For Equal and Double allocation, the average sample size required to achieve the same width of 0.5 using $(1:1)$ and $(2:1)$ randomization techniques have been reported. We see that all the sample sizes, even those obtained using non-RAR techniques, are lesser than that of calculated sample sizes. But our objective of reducing sample size or expected number of cases is attained, as the average sample size for Neyman and RSIHR allocations are lower than that of Equal or Double allocation. Also, the expected sample size for the vaccine group is higher for RSIHR allocation, thereby fulfilling its purpose of reduction of cases. Although, all the RAR designs are successful as per their main objective, ERADE may be considered better than any other as it provides the least sample sizes. The simulated power of the test for VE (at 5% level) have been computed, and compared with its asymptotic power-values, given by (3.4.2), to see if there is any degradation. We see a degradation in power for both RAR and non-RAR designs, but the loss is higher for RAR designs. However, in case of Neyman allocation, ERADE provides similar power to non-RAR designs, and may be considered superior to the other two.

3.7.2 Presence of Drift

Another flaw of the RAR techniques, in general, is that it depends on some assumptions which, if violated, may reflect quite the opposite of what was intended. One instance that may affect the statistical properties of the design is the presence of time trend or drift (Thall et al., 2015). Thall et al. (2015) stated that drift occurs when both p_A and p_B increase by the same amount over time. Drift occurs due to the increase in the standard of health-care over the course of time. They assumed that the patients were entering the trial sequentially, at constant intervals of time, so that the p_A and p_B values change by an equal amount with every entering patient, i.e. for the jth patient,

$$\begin{aligned} p_A(j) &= p_A(j-1) + d \\ &= p_A(0) + jd, \end{aligned}$$

where $p_A(0)$ is the initial value before the start of the trial, and d is the drift term. In other words, the response from patient j is distributed as $Bin(1, p_A(j))$ if he/she is allocated to placebo, and as $Bin(1, p_B(j))$ if allocated to vaccine. Generally, the primary point of interest is in estimating treatment difference $p_B - p_A$, which remains constant over the

course of time, although p_A and p_B vary with time in the presence of drift. They showed that the presence of drift may introduce bias, and reduce the power of the test. We will also incorporate the effect of drift as time trends are quite common during a pandemic; but we will define drift in a slightly different manner. Since, here we are interested in estimating p_B/p_A, rather than the difference $p_B - p_A$, it would be more logical to state that drift occurs if p_B and p_A increase or decrease by same proportion over the same period of time. Again, if we assume that the patients are entering at fixed intervals of time; for the jth patient we have,

$$\begin{aligned} p_A(j) &= p_A(j-1)d \\ &= p_A(0)d^j, j = 1, 2, \ldots \end{aligned}$$

Therefore, the vaccine efficacy remains constant over the course of time, but the individual disease attack-rates change. The assumption is quite legit, as we will see in Chapter 5 that these type of situations are often encountered in time-dependent vaccine efficacy studies, and sometimes the vaccine efficacy may also wane with time. Similarly, as before, if there is a improvement in health-care conditions or seasonal variations, the disease attack rates may decrease, i.e. d becomes less than 1. But, a more practical consideration is the effect of antibody waning in the infected or vaccinated people. As a result, the disease increases in both the vaccinated and unvaccinated cohorts ($d > 1$).

To study the effect of drift in the sample size calculations, the simulations done in the previous section were repeated after incorporating the drift effect, and compared. The expected sample sizes required to achieve a width of 0.5 when trying to target different allocation ratios (Equal, Double, Neyman and RSIHR) using the relevant RAR (SMLE, DBCD, ERADE) or non-RAR ($1:1$, $2:1$ randomization) techniques were calculated in a similar manner as before. However, in this case the jth response in the simulated trial was generated from $Bin(1, p_A(j))$ or $Bin(1, p_B(j))$ (instead of $Bin(1, p_A)$ or $Bin(1, p_B)$), according as the jth individual was allocated to placebo or vaccine. The initial values, $p_A(0)$ and $p_B(0)$ were set equal to the constant p_A and p_B values in the previous section for better comparison. The drift term d was taken to be 1.001, i.e. the attack rates doubled, approximately after every 70 samples. The expected sample sizes (total and vaccine) after considering drift term has been reported in the second part of Table 3.10, and the simulated power has been reported in the last few columns. These sample sizes are not at all comparable to the sample sizes reported in column 2 as those have been calculated without considering the presence of drift. Apparently it may seem from Table 3.10 that drift reduces the sample size, but the real picture is quite different. This happens because d is greater than 1, and with the increase of p_A and p_B, the variance estimate given by (3.2.2), and subsequently the width of the confidence interval, rapidly declines; and the desired width is achieved at a much earlier stage. Also, we see that there has been a slight degradation in power due to the presence of drift. If d is less than 1, then a quite

opposite scenario would have been observed, i.e. underlying p_A and p_B decrease with every incoming patient leading to an increase or fluctuation in the width of the CI. The width may never achieve the desired value, and the expected sample sizes may become infinite. This happens because in the presence of drift, the underlying model changes and the standard estimates do not converge according to the central limit theorem. Therefore, in the presence of drift (even if $d > 1$), the underlying model needs to be updated, and the inference procedure needs to be redefined in light of the new model.

3.8 Concluding Remarks

The sample sizes and variances calculated in this chapter rely on the asymptotic normality assumption for large sample sizes, which hold true for most trials. The main objective was to give an idea about the sample size determination problem in vaccine studies and trace the changes that would occur in the sample size if we deviate from the conventional equal allocation of individuals across the vaccine and placebo, and move towards other optimal allocation schemes. One major disadvantage of the sample size determination problem is that the obtained sample sizes depend on the unknown parameters, VE and ARU which are to be estimated from the trial. To overcome this drawback, we discussed methods for implementing the allocation schemes in an actual trial. But implementation comes with its own flaws, especially when there is drift in the model. However, in the presence of drift we must redefine the model and recalculate the sample sizes as the previous models no longer work. The sequential methods, viz. the SMLE and DBCD procedures, assume that the responses on all the previously recruited individuals are available at the time of recruitment of a new individual. However, that is not always the case in vaccine trials where individuals are assigned to vaccine or placebo and are checked after a specified time for the occurrence of the disease (response). Hence, there might be delayed responses and response on all the previously recruited individuals may not be available at the time of recruitment of a new individual. There is a great deal of literature that incorporates the possibility of delay in RAR designs (Hu et al., 2008, 2009; Bai et al., 2002), most of which suggests working with the available responses only at a time as a common way of handling this delay. However, implementing them in case of vaccine trials where only a few response may be available at the time of new entry require further research.

Moreover, the implementation methods discussed in this chapter help us achieve the desired allocation proportion in the long run but fail to give an estimate of the sample size before the beginning of the trial. Investigators are usually interested in getting an estimate at the initial stage so that the trial can be properly designed. If some prior information is available regarding the values of the unknown parameters, those values can be used to approximate the sample size. The information may be unavailable, and even

without drift								
Allocation Type	Sample Size	Asymptotic Power	Expected Sample Size			Simulated Power		
			SMLE	DBCD	ERADE	SMLE	DBCD	ERADE
Equal	1153 (576)	0.99		1124.18(561.552)			0.945	
Double	1414 (471)	0.99		1080.229(718.153)			0.945	
Neyman	1131 (644)	0.99	1046.129 (581.716)	1052.116 (586.457)	1013.591 (573.534)	0.922	0.929	0.950
RSIHR	1149 (724)	0.99	1024.227 (614.043)	971.199 (593.465)	1011.163 (608.708)	0.925	0.938	0.921
with drift d = 1.001								
Allocation Type	Sample Size	Asymptotic Power	Expected Sample Size			Simulated Power		
			SMLE	DBCD	ERADE	SMLE	DBCD	ERADE
Equal	1153 (576)	-		690.247 (345.049)			0.962	
Double	1414 (471)	-		697.067 (463.77)			0.931	
Neyman	1131 (644)	-	659.63 (370.255)	680.133 (382.316)	662.715 (377.113)	0.942	0.922	0.931
RSIHR	1149 (724)	-	671.526 (403.933)	634.842 (392.628)	674.085 (417.676)	0.917	0.918	0.905

Table 3.10: Table showing the expected sample sizes required to achieve a width 0.5 of the confidence interval for VE under different allocation techniques and randomization (RAR and non-RAR) procedures. The true values of VE and ARU are respectively taken as 0.4 and 0.1, and simulations in both presence and absence of drift have been considered. The simulated and asymptotic powers are also shown, and the corresponding sample sizes for the vaccine group have been reported within parentheses.

if it is, it may be somewhat unreliable leading to uncertain or wrong sample sizes. This impediment can be circumvented if one takes the Bayesian approach described in the next chapter.

Chapter 4

Sample size determination in vaccine trials: A Bayesian Approach

4.1 Introduction

As the title suggests, this chapter considers a Bayesian approach for obtaining the sample sizes in the vaccine trials. It gives a Bayesian alternative to the classical sample size determination problem, described in Chapter 3. While the frequentist approach involved obtaining CIs of VE, the Bayesian approach takes into consideration credible intervals for VE. The Bayesian approach may be particularly attractive to researchers as it can handle the uncertainties associated with the assumed values of the unknown parameters (Adcock, 1987, 1988, 1997). Often, the sample sizes depend on the unknown parameters (VE, ARU, and ARV in the case of comparative cohort vaccine trials) to be estimated from the trial. In such a scenario, one uses the prior information, usually available from some previously conducted trial or from the Phase I/II trial, to guess the underlying values of VE and ARU. At the end of Chapter 3, we pointed out that ambiguities associated with these guesses can provide misleading results, and the frequentist approach has no way to counter such a problem. In contrast, the Bayesian approach can take care of these uncertainties by appropriately choosing the priors on ARV and ARU (see Section 4.5.4 for more details). However, one hurdle may be that the Bayesian sample size calculations require time-consuming simulations, and researchers are often reluctant to use them. For example, the recently developed Sequential Bayesian Designs for COVID-19 trials (Harrell & Lindsell, 2020) used the frequentist approach to sample size estimation as the rapidly developing nature of the trials did not permit such intensive simulation. We address the limitation in this chapter by outlining methods for complexity reduction.

The rest of the chapter is organized as follows. In Section 4.2, we provide a brief literature review on the Bayesian approach to determining sample sizes. Two credible

interval-based criteria, viz., Average Coverage Criterion and Average Length Criterion have been discussed. In Section 4.3, the existing theory has been discussed in the light of vaccine efficacy studies (comparative cohort designs) so that the optimal sample size and allocation proportion can be obtained. In Section 4.4, we outline the various HPD interval computation techniques for obtaining the HPD intervals of the parameter of interest. In Section 4.5, we determine the Bayesian sample sizes under various allocation techniques and compare them. The advantage of the Bayesian approach over the frequentist one has also been pointed out. In Section 4.6, we consider the problem of Bayesian sample size determination in case-control studies. We compare our results in the context of a real-life trial in Section 4.7. Section 4.8 concludes.

4.2 Preliminaries on Bayesian Sample Size Determination

In a general case, suppose $\mathbf{x_n} = (x_1, x_2, \ldots, x_n), \mathbf{x_n} \in \mathcal{X}$ ($\mathcal{X}$ is the sample space), represents a sample of size n, and $\theta \in \Theta$ the parameter of interest. In the frequentist approach, a $100(1-\alpha)\%$ CI of θ is obtained based on the sample of size n. Let $\theta_L(n,\alpha)$ and $\theta_U(n,\alpha)$ respectively denote the lower and upper limits of the $100(1-\alpha)\%$ CI for θ, i.e.

$$Pr(\theta_L(n,\alpha) \leq \theta \leq \theta_U(n,\alpha)) = 1-\alpha.$$

We then fix the maximum width of the CI, i.e. we set

$$\theta_U(n,\alpha) - \theta_L(n,\alpha) \leq l, \tag{4.2.1}$$

where l is the pre-specified width. Then, for a given l and α, the smallest n satisfying (4.2.1) gives the required sample size. In the Bayesian approach, the precision is set by fixing the width of the $100(1-\alpha)\%$ credible interval of θ instead. The Bayesian literature encompasses a large number of credible (HPD/equal-tailed) interval-based criteria that can be used to compute the sample sizes. Two of them are discussed below.

4.2.1 Average Coverage Criterion (ACC)

Suppose $f(\theta \mid \boldsymbol{\eta})$ is the prior density of θ, with $\boldsymbol{\eta}$ being the vector of hyper-parameters. If $f(\mathbf{x_n} \mid \theta)$ is the likelihood of the data, then the marginal distribution of $\mathbf{x_n}$ is given by,

$$f(\mathbf{x_n} \mid n, \boldsymbol{\eta}) = \int_{\Theta} f(\mathbf{x_n} \mid \theta) f(\theta \mid \boldsymbol{\eta}) d\theta,$$

and the posterior distribution of θ given $\mathbf{x_n}$ is given by,

$$f(\theta \mid \mathbf{x_n}, n, \boldsymbol{\eta}) = \frac{f(\mathbf{x_n} \mid \theta) f(\theta \mid \boldsymbol{\eta})}{f(\mathbf{x_n} \mid n, \boldsymbol{\eta})}.$$

If the posterior distribution is unimodal, then the coverage probability of the HPD interval of length l for θ, denoted by $\alpha_l^*(\mathbf{x_n} \mid n, \boldsymbol{\eta})$, must satisfy

$$\int_{a(\mathbf{x_n}|n,\boldsymbol{\eta})}^{a(\mathbf{x_n}|n,\boldsymbol{\eta})+l} f(\theta \mid \mathbf{x_n}, n, \boldsymbol{\eta}) d\theta = \alpha_l^*(\mathbf{x_n} \mid n, \boldsymbol{\eta}),$$

where $a(\mathbf{x_n} \mid n, \boldsymbol{\eta})$ is the lower limit of the HPD interval of length l.

In ACC, we find the smallest n such that the expected coverage probability is at least $1 - \alpha$ (Joseph et al., 1995, 2008). In other words, ACC finds the smallest n satisfying

$$\int_{\mathcal{X}} \alpha_l^*(\mathbf{x_n} \mid n, \boldsymbol{\eta}) f(\mathbf{x_n} \mid n, \boldsymbol{\eta}) d\mathbf{x_n} \geq 1 - \alpha.$$

4.2.2 Average Length Criterion (ALC)

ALC takes quite the reverse approach of ACC, where instead of fixing the length of the HPD interval for θ, one fixes its coverage, and finds the length of the HPD interval of coverage $1 - \alpha$. If $b_L(\mathbf{x_n}, \alpha \mid n, \boldsymbol{\eta})$ and $b_U(\mathbf{x_n}, \alpha \mid n, \boldsymbol{\eta})$ respectively denote the lower and upper limits of the HPD interval with coverage $1 - \alpha$, and $l_{1-\alpha}^*(\mathbf{x_n} \mid n, \boldsymbol{\eta})$ be its length, then

$$\int_{b_L(\mathbf{x_n},\alpha|n,\boldsymbol{\eta})}^{b_U(\mathbf{x_n},\alpha|n,\boldsymbol{\eta})} f(\theta \mid \mathbf{x_n}, n, \boldsymbol{\eta}) d\theta = 1 - \alpha$$
$$\text{and } b_U(\mathbf{x_n}, \alpha \mid n, \boldsymbol{\eta}) - b_L(\mathbf{x_n}, \alpha \mid n, \boldsymbol{\eta}) = l_{1-\alpha}^*(\mathbf{x_n} \mid n, \boldsymbol{\eta})$$

holds.

In ALC, we find the smallest n such that the expected length is at most l, i.e. the minimum n satisfying

$$\int_{\mathcal{X}} l_{1-\alpha}^*(\mathbf{x_n} \mid \mathbf{n}, \boldsymbol{\eta}) f(\mathbf{x_n} \mid n, \boldsymbol{\eta}) d\mathbf{x_n} \leq l$$

is chosen as the appropriate sample size.

4.3 ACC and ALC in the Context of Vaccine Eficacy Studies

Consider a vaccine trial with comparative cohort design that follows model (1.3.1). In the previous section, we reviewed the Bayesian methods for finding sample sizes by considering a single sample of size n. Therefore, determining the quantity n was of interest. But, here, the determination of placebo and vaccine group sample sizes, n_A and n_B, the optimal allocation ratio ρ ($\rho = \frac{n_A}{n_A+n_B}$) are also of interest along with the total sample size n. In simple words, the above can be stated as a two-sample problem. The extension to two-sample case is much simpler in the frequentist approach where the lower and upper limit of the CI for θ becomes function of n_A and n_B, instead of simply n. That is, $\theta_U(n, \alpha)$ and $\theta_L(n, \alpha)$ in equation (4.2.1) are replaced by $\theta_U(n_A, n_B, \alpha)$ and $\theta_L(n_A, n_B, \alpha)$, respectively. The rest of the procedure is similar as before, and we choose n_A and n_B such that

$$\theta_U(n_A, n_B, \alpha) - \theta_L(n_A, n_B, \alpha) \leq l \tag{4.3.1}$$

is satisfied. We observed in Chapter 3, that if we fix the length of the CI for VE, we must have the interval as

$$\widehat{VE} \mp z_{\frac{\alpha}{2}} \sqrt{v_F(\widehat{p}_A, \widehat{p}_B)},$$

where $v_F(p_A, p_B) = \frac{1-p_A}{p_A^3}.\frac{p_B^2}{n_A} + \frac{p_B}{p_A^2}.\frac{1-p_B}{n_B}$, is the asymptotic variance of $\widehat{VE}$ (see equation (3.2.1)). Therefore, equation (4.3.1) in this case becomes

$$2z_{\frac{\alpha}{2}} \sqrt{v_F(p_A, p_B)} \leq l. \tag{4.3.2}$$

In order to find n_A and n_B, the problem can then broadly be divided into the following two cases:

(a) The allocation proportion, ρ, is a fixed quantity. The Equal and Double allocations fall under this category.

(b) ρ is chosen in such way that it is optimal with respect to some criterion. The Neyman and RSIHR allocations obtained in Chapter 3 belong to this class.

The two-sample problem is not so straightforward in the Bayesian approach, and we need slight extensions to the existing theory to address it M'Lan et al. (2006); M'Lan (2002). We consider independent Beta priors, the conjugate priors which is the most natural choice, on p_A and p_B, i.e., we assume

$$\begin{aligned} p_A &\sim Beta(a_A, b_A), \\ \text{and } p_B &\sim Beta(a_B, b_B), \end{aligned} \tag{4.3.3}$$

independently of each other.

The marginal distribution of the data $(x_{An_A}, x_{Bn_B})'$ is the joint distribution of two independent Beta-Binomial distributions, given by,

$$f(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta}) = \binom{n_A}{x_{An_A}} \frac{Beta(a_A + x_{An_A}, n_A + b_A - x_{An_A})}{Beta(a_A, b_A)} \times \binom{n_B}{x_{Bn_B}} \frac{Beta(a_B + x_{Bn_B}, n_B + b_B - x_{Bn_B})}{Beta(a_B, b_B)}, \tag{4.3.4}$$

where $x_{An_A} = 0, 1, 2, \ldots, n_A$, $x_{Bn_B} = 0, 1, 2, \ldots, n_B$ and $\boldsymbol{\eta} = (a_A, b_A, a_B, b_B)$ is the vector of hyper-parameters.

Our parameter of interest is the single-valued quantity VE, and to find ACC and ALC, we have to obtain the HPD interval for VE from its posterior distribution. Analogous to Section 4.2, we find the HPD intervals of fixed length or fixed coverage. The detailed method of finding the HPD interval has been discussed in the next section. For the time being, suppose $a(x_{An_A}, x_{Bn_B}, \boldsymbol{\eta})$ and $\alpha_l^*(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta})$, respectively, denote the lower limit and coverage of the HPD interval of length l for VE. Similarly, let $(b_L(x_{An_A}, x_{Bn_B}, \boldsymbol{\eta}, \alpha), b_U(x_{An_A}, x_{Bn_B}, \boldsymbol{\eta}, \alpha))$ denote the HPD interval for VE with coverage $1-\alpha$, and $l_{1-\alpha}^*(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta})$ be its length. Therefore, the following must hold:

$$\int_{a(x_{An_A}, x_{Bn_B}, \boldsymbol{\eta})}^{a(x_{An_A}, x_{Bn_B}, \boldsymbol{\eta}) + l} f(\theta \mid x_{An_A}, x_{Bn_B}, \boldsymbol{\eta}) d\theta = \alpha_l^*(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta}),$$

$$\int_{b_L(x_{An_A}, x_{Bn_B}, \boldsymbol{\eta}, \alpha)}^{b_U(x_{An_A}, x_{Bn_B}, \boldsymbol{\eta}, \alpha)} f(\theta \mid x_{An_A}, x_{Bn_B}, \boldsymbol{\eta}) d\theta = 1 - \alpha, \text{ and}$$

$$b_U(x_{An_A}, x_{Bn_B}, \boldsymbol{\eta}, \alpha) - b_L(x_{An_A}, x_{Bn_B}, \boldsymbol{\eta}, \alpha) = l_{1-\alpha}^*(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta}),$$

where $f(\cdot | x_{An_A}, x_{Bn_B}, \boldsymbol{\eta})$ is the posterior distribution of VE given $(x_{An_A}, x_{Bn_B})'$. Define,

$$acc(n_A, n_B, \boldsymbol{\eta}) = \sum_{x_{An_A}=0}^{n_A} \sum_{x_{Bn_B}=0}^{n_B} \alpha_l^*(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta}) f(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta}), \text{ and} \tag{4.3.5}$$

$$alc(n_A, n_B, \boldsymbol{\eta}) = \sum_{x_{An_A}=0}^{n_A} \sum_{x_{Bn_B}=0}^{n_B} l_{1-\alpha}^*(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta}) f(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta}). \tag{4.3.6}$$

Note that, $acc(n_A, n_B, \eta)$ and $alc(n_A, n_B, \eta)$ are, respectively, the expected coverage of HPD intervals of length l, and the expected length of HPD intervals with coverage $1-\alpha$.

We see that the above two quantities are functions of n_A and n_B instead of a single sample n. Like the frequentist approach, the problem may henceforth be divided into two

parts based on whether ρ is prefixed or optimally determined.

4.3.1 Sample Size Determination when ρ is Fixed

It should be noted that the proportion allocated to placebo, ρ, and the sample sizes in the two groups n_A and n_B are related as

$$n_A = cn_B, \tag{4.3.7}$$

where $c = \frac{\rho}{1-\rho}$.

When ρ is fixed, equation (4.3.7) reduces $acc(n_A, n_B, \boldsymbol{\eta})$ and $alc(n_A, n_B, \boldsymbol{\eta})$ to $acc(n_A, \boldsymbol{\eta})$ and $alc(n_A, \boldsymbol{\eta})$, respectively. Thus, ACC and ALC, respectively, seeks the minimum n_A such that $acc(n_A, \boldsymbol{\eta}) \geq 1 - \alpha$ and $alc(n_A, \boldsymbol{\eta}) \leq l$. Once the sample size in the placebo group n_A is obtained, the total sample size n, and the sample size in the vaccine group n_B can be obtained using the relation (4.3.7).

Alternatively, if

$$\begin{aligned} \mathcal{N}_1 &= \{(n_A, n_B) : acc(n_A, n_B, \boldsymbol{\eta}) \geq 1 - \alpha\}, \\ \mathcal{N}_2 &= \{(n_A, n_B) : alc(n_A, n_B, \boldsymbol{\eta}) \leq l, \\ \text{and, } \mathcal{N}' &= \{(n_A, n_B) : n_A = \frac{\rho}{1-\rho} n_B\}, \end{aligned}$$

we can say that ACC and ALC respectively seeks the minimum n_A in the set $\mathcal{N}_1 \cap \mathcal{N}'$ and $\mathcal{N}_2 \cap \mathcal{N}'$, respectively.

4.3.2 Sample Size Determination when ρ has to be Optimally Obtained

In this section, we define a relevant objective function $F : [0, \infty]^2 \to \mathbf{R}$, and choose n_A and n_B in such a way that $F(n_A, n_B)$ is optimized at a given level of precision. The choice of $F(n_A, n_B)$ depends on the objective of the trial. For example, optimal allocations developed in Chapter 3 minimizes the total sample size ($F(n_A, n_B) = n_A + n_B$) and the expected number of people developing the disease ($F(n_A, n_B) = n_A p_A + n_B p_B$) respectively, subject to the constraint (4.3.2). They eventually give rise to Neyman and RSIHR allocations whose optimal ρ-values are respectively given by equations (3.3.2) and (3.3.4).

In the Bayesian approach, if one wants to minimize $F(n_A, n_B)$ at a given level of precision, then ACC and ALC, respectively, searches for the pairs $(n_A^{acc}(F), n_B^{acc}(F)) \in \mathcal{N}_1$

and $(n_A^{alc}(F), n_B^{alc}(F)) \in \mathcal{N}_2$, such that

$$(n_A^{acc}(F), n_B^{acc}(F)) = \underset{(n_A,n_B)\in\mathcal{N}_1}{\arg\min}\ \{F(n_A, n_B)\}, \tag{4.3.8}$$

$$\text{and, } (n_A^{alc}(F), n_B^{alc}(F)) = \underset{(n_A,n_B)\in\mathcal{N}_2}{\arg\min}\ \{F(n_A, n_B)\}. \tag{4.3.9}$$

If more than one pair satisfy the equation (4.3.8) or (4.3.9), then any one of them (or their average) can be chosen as the optimal pair. The optimal ρ and sample size n are then obtained as

$$\begin{aligned}
\rho_{acc}(F) &= \frac{n_A^{acc}(F)}{n_A^{acc}(F) + n_B^{acc}(F)}, \\
n^{acc}(F) &= n_A^{acc}(F) + n_B^{acc}(F), \\
\rho_{alc}(F) &= \frac{n_A^{alc}(F)}{n_A^{alc}(F) + n_B^{alc}(F)}, \\
\text{and, } n^{alc}(F) &= n_A^{alc}(F) + n_B^{alc}(F).
\end{aligned} \tag{4.3.10}$$

To obtain an allocation analogous to that of Neyman in the Bayesian approach, we put $F(n_A, n_B) = n_A + n_B$ in equations (4.3.8) and (4.3.9). For obvious reasons, this allocation is referred to as 'Bayesian Neyman allocation'. Again, an RSIHR like allocation in the Bayesian case minimizes

$$E(x_{An_A} + x_{Bn_B}) = n_A \frac{a_A}{a_A + b_A} + n_B \frac{a_B}{a_B + b_B},$$

over the sets $\mathcal{N}_1$ and $\mathcal{N}_2$. Note that, the expectation of $x_{An_A} + x_{Bn_B}$ (total number of people developing the disease) is taken over the marginal distribution given by (4.3.4).

It should be noted that optimal n will be a function of length l, coverage probability $1-\alpha$, and the hyper-parameters $\boldsymbol{\eta}$. In the Bayesian approach, the precision is specified by fixing the quantity l, and for fixed l and α, one can find the optimal ρ in terms of the hyper-parameters. The approximate formulas for ρ and n have been derived in Section 4.5.

4.4 HPD Interval Computation

It is clear from Sections 4.2 and 4.3 that the first step towards Bayesian sample size determination consists of finding an HPD interval of given length or coverage for VE. This section will focus on different approaches for finding the same. A natural analogy with frequentist case may prompt one to start from an HPD interval of β, viz. (β_1, β_2). However, it is to be noted that even though $(1 - e^{\beta_2}, 1 - e^{\beta_1})$ represent a credible interval

for VE of the same coverage, it may not represent the HPD interval. This is in contrast to the frequentist approach where $100(1-\alpha)\%$ CI for VE could be recovered from a CI for β of the same coverage. But, one case can easily work with ψ instead of VE, as they are linearly related. To find the HPD interval for ψ, we first have to derive its posterior distribution. Even though the exact posterior distribution of ψ can be obtained, the length or coverage of its HPD interval would be difficult to obtain. We consider two approaches to circumvent the problems. The first approach is often used in literature, and it applies Monte Carlo techniques for approximating the HPD intervals (M'Lan et al., 2006; Joseph et al., 2008). The second approach ignores the exact distribution altogether, and works with the asymptotic distribution of ψ. In the next part, we discuss the two approaches. The form of the exact posterior distribution, and the difficulties in working with this form have also been discussed.

4.4.1 Exact Posterior Distribution of ψ

The posterior distribution of p_A and p_B given the data (x_{An_A}, x_{Bn_B}) are independent $Beta(a_A + x_{An_A}, b_A + n_A - x_{An_A})$ and $Beta(a_B + x_{Bn_B}, b_B + n_B - x_{Bn_B})$ respectively. Therefore, the posterior density of $\psi = \frac{p_B}{p_A}$ is given by,

$$g(\psi \mid x_{An_A}, x_{Bn_B}, \boldsymbol{\eta}) = \begin{cases} \frac{B(a_A^*+a_B^*,b_A^*)}{B(a_A^*,b_A^*)B(a_B^*,b_B^*)}\psi^{a_B^*-1} \\ \times {}_2F_1(a_A^*+a_B^*, 1-b_B^*, b_A^*+a_A^*+a_B^*; \psi) \quad \text{if } 0<\psi<1, \\ \frac{B(a_A^*+a_B^*,b_B^*)}{B(a_A^*,b_A^*)B(a_B^*,b_B^*)}\psi^{a_A^*-1} \\ \times {}_2F_1(a_A^*+a_B^*, 1-b_A^*, b_B^*+a_A^*+a_B^*; \frac{1}{\psi}) \quad \text{if } \psi \geq 1, \end{cases} \tag{4.4.1}$$

where ${}_2F_1(a,b,c;x) = \int_0^1 \frac{u^{a-1}(1-u)^{c-a-1}(1-xu)^{-b}}{B(a,c-a)}du$ is the Gauss hypergeometric function in 3 parameters, with $a_A^* = a_A + x_{An_A}$, $b_A^* = b_A + n_A - x_{An_A}$, $a_B^* = a_B + x_{Bn_B}$ and $b_B^* = b_B + n_B - x_{Bn_B}$ (Pham-Gia, 2000).

The density g is unimodal (M'Lan, 2002), and any $100(1-\alpha)\%$ HPD interval for ψ, (ψ_1, ψ_2) must satisfy,

$$g(\psi_1 \mid x_{An_A}, x_{Bn_B}, \boldsymbol{\eta}) = g(\psi_2 \mid x_{An_A}, x_{Bn_B}, \boldsymbol{\eta}) \text{ and} \tag{4.4.2}$$

$$G(\psi_2 \mid x_{An_A}, x_{Bn_B}, \boldsymbol{\eta})) - G(\psi_1 \mid x_{An_A}, x_{Bn_B}, \boldsymbol{\eta}) = 1-\alpha, \tag{4.4.3}$$

where $G(.)$ is the CDF of g.

Therefore, if ψ_1 and ψ_2 are solutions to the set of equations (4.4.2) and (4.4.3), then

$$l_{1-\alpha}^*(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta}) = \psi_2 - \psi_1.$$

Conversely, the coverage of the HPD interval of length l viz. $(\psi^*, \psi^* + l)$ can be obtained as

$$\alpha_l^*(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta}) = \int_{\psi^*}^{\psi^*+l} g(\psi \mid x_{An_A}, x_{Bn_B}, \boldsymbol{\eta}) d\psi,$$

where ψ^* is the solution to

$$g(\psi^* \mid x_{An_A}, x_{Bn_B}, \boldsymbol{\eta}) = g(\psi^* + l \mid x_{An_A}, x_{Bn_B}, \boldsymbol{\eta}). \tag{4.4.4}$$

It should be noted that due to the presence of the Gauss hypergeometric function, equations (4.4.2), (4.4.3) and (4.4.4) are in mathematically intractable forms. It would be difficult to obtain the solutions even after applying numerical methods. Hence, we try to approximate the length and coverage of the HPD intervals using Monte Carlo methods.

4.4.2 Monte Carlo Techniques for Approximating HPD Intervals

As the posterior distribution of ψ is unimodal, Monte Carlo methods may well be applied to estimate the HPD interval of ψ. The Monte Carlo methods consist of generating N independent observations from the (unimodal) posterior distribution of the concerned parameter (ψ in our case). For given (x_{An_A}, x_{Bn_B}), it is quite easy to generate N independent samples, viz. $p_{A1}, p_{A2}, \ldots, p_{AN}$ and $p_{B1}, p_{B2}, \ldots, p_{BN}$, from the posterior distributions of p_A and p_B. Note that,

$$\psi_i = \frac{p_{Bi}}{p_{Ai}}, \ i = 1, 2, \ldots, N$$

represents N independent samples simulated from the posterior distribution of $\psi | x_{An_A}, x_{Bn_B}$.

To estimate the HPD interval of a given length l, the coverages of the set of all credible intervals of length l, $(\psi_{(i)}, \psi_{(i)} + l), i = 1, 2, \ldots, N$, were computed. The interval with the highest coverage was taken as an estimate of the HPD interval for ψ, and its coverage was an estimate of $\alpha_l^*(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta})$ (Chen & Shao, 1999).

Therefore, we have,

$$\widehat{\alpha}_l^*(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta}) = \max_{1 \le i \le N} \frac{\#\{1 \le j \le N : \psi_{(i)} \le \psi_j \le \psi_{(i)} + l\}}{N}. \tag{4.4.5}$$

Similarly, the set of all Monte-Carlo credible intervals of coverage $1 - \alpha$ were considered, and the one with the smallest length was chosen as the estimate of the $100(1-\alpha)\%$ HPD interval of ψ (see Joseph *et al.*, 2008). The length of the HPD interval $l_{1-\alpha}^*(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta})$, $(\psi_{(i)}, \psi_{(i+[(1-\alpha)N])})$, $i = 1, 2, \ldots, N - [(1-\alpha)N]$, can then be estimated as

$$\widehat{l}_{1-\alpha}^*(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta}) = \min_{1 \le i \le N-[(1-\alpha)N]} (\psi_{(i+[(1-\alpha)N])} - \psi_{(i)}). \tag{4.4.6}$$

The estimates $\widehat{\alpha}_l^*(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta})$ and $\widehat{l}_{1-\alpha}^*(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta})$ are then plugged in to equations (4.3.5) and (4.3.6) to get $\widehat{acc}(n_A, n_B, \eta)$ and $\widehat{alc}(n_A, n_B, \eta)$, respectively. However, the computation of $\widehat{l}_{1-\alpha}^*(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta})$ and $\widehat{\alpha}_l^*(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta})$ for every possible combination of x_{An_A} and x_{Bn_B} makes the Monte Carlo procedure computationally cumbersome and time consuming.

4.4.3 Asymptotic Posterior Distribution of ψ

The intractability of the exact distribution of ψ, combined with the time and computation complexity of the Monte Carlo methods motivates us to rely on its asymptotic form to find the HPD intervals. The asymptotic posterior distribution of ψ has been derived in Appendix C.1 using the Bernstein Von-Mises Theorem (Ghosh et al., 2006; Van der Vaart, 1998), and its form is given by

$$\psi \mid x_{An_A}, x_{Bn_B} \overset{a}{\sim} N\left(\widetilde{\psi}, v(\widetilde{p_A}, \widetilde{p_B})\right), \tag{4.4.7}$$

where $\widetilde{p_A} = \frac{a_A + x_{An_A}}{a_A + b_A + n_A - 1}$, $\widetilde{p_B} = \frac{a_B + x_{Bn_B} - 1}{a_B + b_B + n_B - 2}$, $\widetilde{\psi} = \frac{\widetilde{p_B}}{\widetilde{p_A}}$, and $v(\widetilde{p_A}, \widetilde{p_B}) = \frac{\widetilde{p_B}^2(1-\widetilde{p_A})}{\widetilde{p_A}^3(a_A + b_A + n_A - 1)} + \frac{\widetilde{p_B}(1-\widetilde{p_B})}{\widetilde{p_A}^2(a_B + b_B + n_B - 2)}$. Approximating the posterior distribution of ψ by (4.4.7), we get the length and coverage of the HPD interval as,

$$l_{1-\alpha}^*(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta}) = 2z_{\frac{\alpha}{2}}\sqrt{v(\widetilde{p_A}, \widetilde{p_B})}$$

$$\text{and } \alpha_l^*(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta}) = 2\Phi\left(\frac{l}{2\sqrt{v(\widetilde{p_A}, \widetilde{p_B})}}\right) - 1.$$

Therefore, $alc(n_A, n_B, \boldsymbol{\eta})$ and $acc(n_A, n_B, \boldsymbol{\eta})$ can explicitly be written as

$$\begin{aligned} acc(n_A, n_B, \eta) &= \sum_{x_{An_A}=0}^{n_A} \sum_{x_{Bn_B}=0}^{n_B} \left\{2\Phi\left(\frac{l}{2\sqrt{v(\widetilde{p_A}, \widetilde{p_B})}}\right) - 1\right\} \\ &\qquad \times f(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta}) \\ &= E_{(x_{An_A}, x_{Bn_B})}\left(2\Phi\left(\frac{l}{2\sqrt{v(\widetilde{p_A}, \widetilde{p_B})}}\right) - 1\right), \end{aligned} \tag{4.4.8}$$

$$alc(n_A, n_B, \eta) = E_{(x_{An_A}, x_{Bn_B})}\left(2z_{\frac{\alpha}{2}}\sqrt{v(\widetilde{p_A}, \widetilde{p_B})}\right). \tag{4.4.9}$$

We see that $acc(n_A, n_B, \boldsymbol{\eta})$ and $alc(n_A, n_B, \boldsymbol{\eta})$ will be functions of the hyper-parameters $\boldsymbol{\eta}$ and sample sizes n_A,n_B. But, obtaining closed form expression for them is difficult.

4.5 Determination of Optimal Allocation Proportion and Sample Size

The HPD intervals, obtained using different approaches mentioned in Section 4.4, can be used to calculate the optimal allocation ratio and sample sizes, following the procedure mentioned in Section 4.3. Since $acc(n_A, n_B, \boldsymbol{\eta})$, $alc(n_A, n_B, \boldsymbol{\eta})$ and subsequently the sample sizes obtained from them vary according to the HPD interval computation technique used, we obtain the sample sizes (under both fixed and optimal allocation techniques) using all possible combinations of HPD interval computation methods and criterion functions (ACC, ALC). In the absence of closed form expressions for $acc(n_A, n_B, \boldsymbol{\eta})$ and $alc(n_A, n_B, \boldsymbol{\eta})$, we fail to get an explicit formula for the sample size as in the frequentist case (see equations (3.2.5) and (3.2.6)). However, certain approximate bounds (closed form) can be obtained in case of Bayesian sample sizes, which can then be used to get an idea about the sample size. In the next part, we obtain the bounds on the fixed and the two optimal allocations, Neyman and RSIHR.

4.5.1 Fixed Allocation Technique

If the allocation ratio ρ is fixed, $acc(n_A, n_B, \boldsymbol{\eta})$ and $alc(n_A, n_B, \boldsymbol{\eta})$ simply reduce to $acc(n_A, \boldsymbol{\eta})$ and $alc(n_A, \boldsymbol{\eta})$. Therefore, in order to obtain the ACC/ALC sample sizes, we start by computing $acc(n_A, \boldsymbol{\eta})/\ alc(n_A, \boldsymbol{\eta})$ numerically for $n_A = 1$, and gradually increase the values of n_A to find the point where it crosses/becomes less than $(1-\alpha)/\ l$. The re-computation of $acc(n_A, \boldsymbol{\eta})$ and $alc(n_A, \boldsymbol{\eta})$ several times make the procedure computationally intensive and time consuming, especially when we follow the Monte Carlo approach for HPD interval computation. However, the computational complexity involved in evaluating $acc(n_A, n_B, \boldsymbol{\eta})$ and $alc(n_A, n_B, \boldsymbol{\eta})$ can be reduced considerably by replacing $\sqrt{v(\widetilde{p_A}, \widetilde{p_B})}$ by $E_{\boldsymbol{x}}(\sqrt{v(\widetilde{p_A}, \widetilde{p_B})})$ or $\sqrt{E_{\boldsymbol{x}}(v(\widetilde{p_A}, \widetilde{p_B}))}$, and thereby considering

$$acc'(n_A, n_B, \boldsymbol{\eta}) = 2\Phi\left(\frac{l}{2E_{\boldsymbol{x}}(\sqrt{v(\widetilde{p_A}, \widetilde{p_B})})}\right) - 1,$$

$$\text{or } acc''(n_A, n_B, \boldsymbol{\eta}) = 2\Phi\left(\frac{l}{2\sqrt{E_{\boldsymbol{x}}(v(\widetilde{p_A}, \widetilde{p_B}))}}\right) - 1,$$

instead of $acc(n_A, n_B, \boldsymbol{\eta})$ in Bayesian sample size determination. The sample sizes so obtained will be henceforth referred to as ACC' and ACC'' sample sizes. In literature, the unknowns in the variance term have often been replaced by their prior expectations (Adcock, 1987, 1988). The advantage of the substitution is that we can obtain some upper bounds on the resultant ACC', ACC'' and even ALC sample sizes. The upper bounds

are much easier to calculate, and give us an approximate idea about the sample size. In fact, it has been shown that ACC' and ALC sample sizes are the same.

We begin with defining the following quantities:

$$\begin{aligned}
alc'(n_A, n_B, \boldsymbol{\eta}) &= 2z_{\frac{\alpha}{2}}\sqrt{E_{\boldsymbol{x}}(v(\widetilde{p_A}, \widetilde{p_B}))},\\
\mathcal{N}_1' &= \{(n_A, n_B) : acc'(n_A, n_B, \boldsymbol{\eta}) \geq 1-\alpha\},\\
\mathcal{N}_1'' &= \{(n_A, n_B) : acc''(n_A, n_B, \boldsymbol{\eta}) \geq 1-\alpha\},\\
\mathcal{N}_2' &= \{(n_A, n_B) : alc'(n_A, n_B, \boldsymbol{\eta}) \leq l\},\\
n_A^{acc}(f) &= \min_{(n_A,n_B)\in\mathcal{N}_1\cap\mathcal{N}'} n_A,\\
n_A^{acc'}(f) &= \min_{(n_A,n_B)\in\mathcal{N}_1'\cap\mathcal{N}'} n_A,\\
n_A^{acc''}(f) &= \min_{(n_A,n_B)\in\mathcal{N}_1''\cap\mathcal{N}'} n_A,\\
\text{and, } n_A^{alc'}(f) &= \min_{(n_A,n_B)\in\mathcal{N}_2'\cap\mathcal{N}'} n_A.
\end{aligned}$$

Note that, $n_A^{acc}(f)$ is the actual ACC sample size, whereas $n_A^{acc'}(f)$ and $n_A^{acc''}(f)$ are respectively the ACC' and ACC'' sample sizes. Before mentioning the main result to obtain the bounds, we state a relevant lemma.

Lemma 4.5.1. *If the hyper-parameters a_A, b_A, a_B and b_B are such that $a_A > 3$, $b_A > 1$, $a_B > 1$ and $b_B > 1$, then*

$$\begin{aligned}
E_{\boldsymbol{x}}(v(\widetilde{p_A}, \widetilde{p_B})) \leq\ & \frac{(a_B+b_B+n_B)(a_B+b_B+n_B+1)}{(a_A+b_A+n_A-2)(a_B+b_B+n_B-2)^2}.v_1(\boldsymbol{\eta})\\
&+ \frac{(a_A+b_A+n_A-1)(a_B+b_B+n_B)(a_B+b_B+n_B+1)}{(a_A+b_A+n_A-2)(a_B+b_B+n_B-2)^3}.v_2(\boldsymbol{\eta}), \qquad (4.5.1)
\end{aligned}$$

where $v_1(\boldsymbol{\eta}) = \frac{(a_A+b_A-1)(a_A+b_A-2)b_A}{(a_A-1)(a_A-2)(a_A-3)}.\frac{a_B(a_B+1)}{(a_B+b_B)(a_B+b_B+1)}$ *and*
$v_2(\boldsymbol{\eta}) = \frac{(a_A+b_A-1)(a_A+b_A-2)}{(a_A-1)(a_A-2)}.\frac{a_Bb_B}{(a_B+b_B)(a_B+b_B+1)}$.

Further, if n_A and n_B are such the $n_A \longrightarrow \infty$, $n_B \longrightarrow \infty$ and $\frac{n_A}{n_A+n_B} \longrightarrow \rho$, then

$$E_{\boldsymbol{x}}(v(\widetilde{p_A}, \widetilde{p_B})) \lesssim \frac{v_1(\boldsymbol{\eta})}{(a_A+b_A+n_A-2)} + \frac{v_2(\boldsymbol{\eta})}{(a_B+b_B+n_B-2)}. \qquad (4.5.2)$$

Proof. See Appendix C.2 □

Let us denote the quantity on the RHS of (4.5.1) by $v(\boldsymbol{\eta}, n_A, n_B)$. Again define,

$$
\begin{aligned}
&acc'''(n_A, n_B, \boldsymbol{\eta}) = 2\Phi\left(\frac{l}{2\sqrt{v(\boldsymbol{\eta}, n_A, n_B)}}\right) - 1,\\
&alc''(n_A, n_B, \boldsymbol{\eta}) = 2z_{\frac{\alpha}{2}}\sqrt{v(\boldsymbol{\eta}, n_A, n_B)},\\
&\mathcal{N}_1''' = \{(n_A, n_B) : acc'''(n_A, n_B, \boldsymbol{\eta}) \geq 1-\alpha\},\\
&\mathcal{N}_2'' = \{(n_A, n_B) : alc''(n_A, n_B, \boldsymbol{\eta}) \leq l\},\\
&n_A^{acc'''}(f) = \min_{(n_A,n_B)\in\mathcal{N}_1'''\cap\mathcal{N}'} n_A\\
\text{and, } &n_A^{alc''}(f) = \min_{(n_A,n_B)\in\mathcal{N}_2''\cap\mathcal{N}'} n_A.
\end{aligned}
$$

Result 4.5.1. *If the hyper-parameters satisfy the conditions of Lemma 4.5.1, then*

(i) $n_A^{acc'}(f) \leq n_A^{acc''}(f) \leq n_A^{acc'''}(f)$

(ii) $n_A^{alc}(f) \leq n_A^{alc'}(f) \leq n_A^{alc''}(f)$

(iii) Further, if the hyper-parameters are such that the quadratic equation

$$
\frac{v_1(\boldsymbol{\eta})}{a_A + b_A - 2 + x} + \frac{cv_2(\boldsymbol{\eta})}{c(a_B + b_B - 2) + x} = \frac{1}{4}\left(\frac{l}{z_{\frac{\alpha}{2}}}\right)^2 \tag{4.5.3}
$$

has at least one positive root, greater than or equal to 1, then

$$
n_A^{acc'''}(f) = n_A^{alc''}(f) \approx n_A^{sol}(\boldsymbol{\eta}, \alpha, l, c),
$$

where $n_A^{sol}(\boldsymbol{\eta}, \alpha, l, c)$ *is the maximum root of equation (4.5.3).*

Proof. (i) Using Jensen's equality and Lemma 4.5.1, we have,

$$
\begin{aligned}
&E_{\boldsymbol{x}}(\sqrt{v(\widetilde{p_A}, \widetilde{p_B})}) \leq \sqrt{E_{\boldsymbol{x}}(v(\widetilde{p_A}, \widetilde{p_B}))} \leq \sqrt{v(\boldsymbol{\eta}, n_A, n_B)}\\
\implies &acc'(n_A, n_B, \boldsymbol{\eta}) \geq acc''(n_A, n_B, \boldsymbol{\eta}) \geq acc'''(n_A, n_B, \boldsymbol{\eta})\\
\implies &\mathcal{N}_1' \supseteq \mathcal{N}_1'' \supseteq \mathcal{N}_1'''\\
\implies &n_A^{acc'}(f) \leq n_A^{acc''}(f) \leq n_A^{acc'''}(f).
\end{aligned}
$$

(ii) It is easy to see that,

$$
\begin{aligned}
&\mathcal{N}_1' = \mathcal{N}_2,\ \mathcal{N}_1'' = \mathcal{N}_2',\ \text{and } \mathcal{N}_1''' = \mathcal{N}_2''\\
\implies &n_A^{acc'}(f) = n_A^{alc}(f),\ n_A^{acc''}(f) = n_A^{alc'}(f),\ \text{and } n_A^{acc'''}(f) = n_A^{alc''}(f))\\
\implies &n_A^{alc}(f) \leq n_A^{alc'}(f) \leq n_A^{alc''}(f).
\end{aligned}
$$

It is also evident that the ALC and ALC' sample sizes are respectively equal to the ACC' and ACC'' sample sizes.

(iii) From the second part of Lemma 4.5.1, we have,

$$v(\boldsymbol{\eta}, n_A, n_B) \approx \frac{v_1(\boldsymbol{\eta})}{(a_A + b_A + n_A - 2)} + \frac{v_2(\boldsymbol{\eta})}{(a_B + b_B + n_B - 2)}.$$

Note that,

$$\begin{aligned}
&(n_A, n_B) \in \mathcal{N}_1''' \cap \mathcal{N} \\
\implies &acc'''(n_A, n_B, \boldsymbol{\eta}) \geq (1-\alpha) \text{ and } n_A = cn_B \\
\implies &v(\boldsymbol{\eta}, n_A, n_A/c) \leq \frac{1}{4}\left(\frac{l}{z_{\frac{\alpha}{2}}}\right)^2 \\
\implies &\frac{v_1(\boldsymbol{\eta})}{(a_A + b_A + n_A - 2)} + \frac{v_2(\boldsymbol{\eta})}{(a_B + b_B + n_B - 2)} \lesssim \frac{1}{4}\left(\frac{l}{z_{\frac{\alpha}{2}}}\right)^2 \text{ (using (4.5.2))} \\
\implies &n_A \gtrsim n_A^{sol}(\boldsymbol{\eta}, \alpha, l, c).
\end{aligned}$$

Hence, the proof is complete. □

From the above result, it can be seen that ACC', ACC'' and ALC sample sizes will always be less than the upper bound $n_A^{sol}(\boldsymbol{\eta}, \alpha, l, c)$ for large n_A and n_B. Therefore, $n_A^{sol}(\boldsymbol{\eta}, \alpha, l, c)$ and $n_B^{sol}(\boldsymbol{\eta}, \alpha, l, c) = \frac{n_A^{sol}(\boldsymbol{\eta}, \alpha, l, c)}{c}$ give us an idea about the approximate sample size to be taken for the study. The bounds can be useful even if one is interested in numerically determining the actual ACC and ALC sample sizes. They can serve as seed values, i.e. instead of starting with $n_A = 1$ and gradually increasing the sample sizes, one can start with $n_A = n_A^{sol}(\boldsymbol{\eta}, \alpha, l, c)$ and search around it to get the actual sample size with much lesser computational effort.

4.5.2 Bayesian Neyman Allocation

As mentioned in Section 4.3.2, ACC and ALC, respectively searches for the pairs $(n_A^{acc}(N), n_B^{acc}(N)) \in \mathcal{N}_1$ and $(n_A^{alc}(N), n_B^{alc}(N)) \in \mathcal{N}_2$, such that

$$\begin{aligned}
(n_A^{acc}(N), n_B^{acc}(N)) &= \underset{(n_A, n_B) \in \mathcal{N}_1}{\arg\min} \{n_A + n_B\}, \\
\text{and } (n_A^{alc}(N), n_B^{alc}(N)) &= \underset{(n_A, n_B) \in \mathcal{N}_2}{\arg\min} \{n_A + n_B\}.
\end{aligned} \tag{4.5.4}$$

It is to be noted that the optimal allocation techniques have a two fold increase in computational complexity as compared to fixed allocation, and applying numerical methods

(especially Monte Carlo techniques) becomes extremely time-consuming. However, approximate bounds on the total sample size can be obtained. The bounds can either be directly used to get an idea about the sample size or be fed into the machine as a starting value and a grid search around it can give us the actual sample size without much computational burden.

Before obtaining the bounds, we define $(n_A^{acc'}(N), n_B^{acc'}(N))$, $(n_A^{acc''}(N), n_B^{acc''}(N))$, $(n_A^{acc'''}(N), n_B^{acc'''}(N))$, $(n_A^{alc'}(N), n_B^{alc'}(N))$ and $(n_A^{alc''}(N), n_B^{alc''}(N))$ by replacing $\mathcal{N}_1$ in equation (4.5.4) with $\mathcal{N}_1'$, $\mathcal{N}_1''$, $\mathcal{N}_1'''$, $\mathcal{N}_2'$ and $\mathcal{N}_2''$, respectively. Like (4.3.10), $\rho_{acc}(N)$, $\rho_{acc'}(N)$, $\rho_{acc''}(N)$, $\rho_{acc'''}(N)$, $\rho_{alc}(N)$ and $\rho_{alc'}(N)$ can also be defined corresponding to each of the above pairs.

Result 4.5.2. *Under the conditions on the hyper-parameters stated in Lemma 4.5.1,*

(i) $n_A^{acc'}(N) + n_B^{acc'}(N) \leq n_A^{acc''}(N) + n_B^{acc''}(N) \leq n_A^{acc'''}(N) + n_B^{acc'''}(N)$.

(ii) $n_A^{alc}(N) + n_B^{alc}(N) \leq n_A^{alc'}(N) + n_B^{alc'}(N) \leq n_A^{alc''}(N) + n_B^{alc''}(N)$.

(iii)
$$n_A^{acc'''}(N) + n_B^{acc'''}(N) = n_A^{alc''}(N) + n_B^{alc''}(N)$$
$$\approx 4\left(\frac{z_{\frac{\alpha}{2}}}{l}\right)^2 (\sqrt{v_1(\boldsymbol{\eta})} + \sqrt{v_2(\boldsymbol{\eta})})^2 - (a_A + b_A + a_B + b_B - 4)$$
$$\rho_{acc'''}(N) = \rho_{alc''}(N) \approx \frac{4\left(\frac{z_{\frac{\alpha}{2}}}{l}\right)^2 \sqrt{v_1(\boldsymbol{\eta})}(\sqrt{v_1(\boldsymbol{\eta})} + \sqrt{v_2(\boldsymbol{\eta})}) - (a_A + b_A - 2)}{4\left(\frac{z_{\frac{\alpha}{2}}}{l}\right)^2 (\sqrt{v_1(\boldsymbol{\eta})} + \sqrt{v_2(\boldsymbol{\eta})})^2 - (a_A + b_A + a_B + b_B - 4)}$$

Proof. (i) It has already been shown that,

$$\mathcal{N}_1' \supseteq \mathcal{N}_1'' \supseteq \mathcal{N}_1'''$$
$$\implies n_A^{acc'}(N) + n_B^{acc'}(N) \leq n_A^{acc''}(N) + n_B^{acc''}(N) \leq n_A^{acc'''}(N) + n_B^{acc'''}(N)$$

(ii) This part is obvious since $\mathcal{N}_1' = \mathcal{N}_2$, $\mathcal{N}_1'' = \mathcal{N}_2'$, and $\mathcal{N}_1''' = \mathcal{N}_2''$.

(iii) Again,

$$acc'''(n_A, n_B, \boldsymbol{\eta}) \geq (1 - \alpha)$$
$$\implies v(\boldsymbol{\eta}, n_A, n_B) \leq \frac{1}{4}\left(\frac{l}{z_{\frac{\alpha}{2}}}\right)^2$$
$$\implies \frac{v_1(\boldsymbol{\eta})}{(a_A + b_A + n_A - 2)} + \frac{v_2(\boldsymbol{\eta})}{(a_B + b_B + n_B - 2)} \lesssim \frac{1}{4}\left(\frac{l}{z_{\frac{\alpha}{2}}}\right)^2 \text{ (using (4.5.2))}$$

Hence, $\mathcal{N}_1''' \approx \left\{(n_A, n_B) : \frac{v_1(\boldsymbol{\eta})}{(a_A + b_A + n_A - 2)} + \frac{v_2(\boldsymbol{\eta})}{(a_B + b_B + n_B - 2)} \leq \frac{1}{4}\left(\frac{l}{z_{\frac{\alpha}{2}}}\right)^2\right\}$.

To obtain $n_A^{acc'''}(N)$ and $n_B^{acc'''}(N)$, we have to minimize $n_A + n_B$ over $\mathcal{N}_1'''$, which is roughly equivalent to solving the constrained optimization problem -

$$\min\ n_A + n_B \text{ such that}$$
$$\frac{v_1(\boldsymbol{\eta})}{(a_A + b_A + n_A - 2)} + \frac{v_2(\boldsymbol{\eta})}{(a_B + b_B + n_B - 2)} \leq \frac{1}{4}\left(\frac{l}{z_{\frac{\alpha}{2}}}\right)^2.$$

Hence, $n_A^{acc'''}(N)$ and $n_B^{acc'''}(N)$ are obtained as

$$n_A^{acc'''}(N) \approx 4\left(\frac{z_{\frac{\alpha}{2}}}{l}\right)^2 \sqrt{v_1(\boldsymbol{\eta})}(\sqrt{v_1(\boldsymbol{\eta})} + \sqrt{v_2(\boldsymbol{\eta})}) - (a_A + b_A - 2)$$
$$\text{and, } n_B^{acc'''}(N) \approx 4\left(\frac{z_{\frac{\alpha}{2}}}{l}\right)^2 \sqrt{v_2(\boldsymbol{\eta})}(\sqrt{v_1(\boldsymbol{\eta})} + \sqrt{v_2(\boldsymbol{\eta})}) - (a_B + b_B - 2).$$

The rest of the proof is immediate.

□

4.5.3 Bayesian RSIHR Allocation

In Bayesian RSIHR allocation, ACC and ALC, respectively, searches for the pairs $(n_A^{acc}(R), n_B^{acc}(R)) \in \mathcal{N}_1$ and $(n_A^{alc}(R), n_B^{alc}(R)) \in \mathcal{N}_2$ such that

$$(n_A^{acc}(R), n_B^{acc}(R)) = \underset{(n_A,n_B)\in\mathcal{N}_1}{\arg\min}\ \{n_A\frac{a_A}{a_A + b_A} + n_B\frac{a_B}{a_B + b_B}\}, \tag{4.5.5}$$
$$\text{and, } (n_A^{alc}(R), n_B^{alc}(R)) = \underset{(n_A,n_B)\in\mathcal{N}_2}{\arg\min}\ \{n_A\frac{a_A}{a_A + b_A} + n_B\frac{a_B}{a_B + b_B}\}.$$

The corresponding allocation proportions are then obtained as,

$$\rho_{acc}(R) = \frac{n_A^{acc}(R)}{n_A^{acc}(R) + n_B^{acc}(R)}$$
$$\text{and, } \rho_{alc}(R) = \frac{n_A^{alc}(R)}{n_A^{alc}(R) + n_B^{alc}(R)}.$$

Evidently, $(n_A^{acc'}(R), n_B^{acc'}(R))$, $(n_A^{acc''}(R), n_B^{acc''}(R))$, $(n_A^{acc'''}(R), n_B^{acc'''}(R))$, $(n_A^{acc'''}(R)$, $n_B^{acc'''}(R))$, $(n_A^{alc'}(R), n_B^{alc'}(R))$ and $(n_A^{alc''}(R), n_B^{alc''}(R))$ can be defined by replacing $\mathcal{N}_1$ in equation (4.5.5) by $\mathcal{N}_1'$, $\mathcal{N}_1''$, $\mathcal{N}_1'''$, $\mathcal{N}_2'$ and $\mathcal{N}_2''$, respectively, and we can state a result similar to Result 4.5.2 in this case. The pertinent result is stated below and its proof has been omitted as it is similar to the proof of Result 4.5.2.

Result 4.5.3. *It can be shown that*

$$(i)\ n_A^{acc'}(R)\frac{a_A}{a_A+b_A}+n_B^{acc'}(R)\frac{a_B}{a_B+b_B}\leq n_A^{acc''}(R)\frac{a_A}{a_A+b_A}+n_B^{acc''}(R)\frac{a_B}{a_B+b_B}$$
$$\leq n_A^{acc'''}(R)\frac{a_A}{a_A+b_A}+n_B^{acc'''}(R)\frac{a_B}{a_B+b_B}.$$

$$(ii)\ n_A^{alc}(R)\frac{a_A}{a_A+b_A}+n_B^{alc}(R)\frac{a_B}{a_B+b_B}\leq n_A^{alc'}(R)\frac{a_A}{a_A+b_A}+n_B^{alc'}(R)\frac{a_B}{a_B+b_B}$$
$$\leq n_A^{alc''}(R)\frac{a_A}{a_A+b_A}+n_B^{alc''}(R)\frac{a_B}{a_B+b_B}.$$

$$(iii)\ n_A^{acc'''}(R)=n_A^{alc''}(R)$$
$$\approx 4\left(\frac{z_{\frac{\alpha}{2}}}{l}\right)^2(\sqrt{v_1(\boldsymbol{\eta})}\sqrt{\frac{a_A}{a_A+b_A}}+\sqrt{v_2(\boldsymbol{\eta})}\sqrt{\frac{a_B}{a_B+b_B}})\times$$
$$\sqrt{v_1(\boldsymbol{\eta})}\sqrt{\frac{a_A+b_A}{a_A}}-(a_A+b_A-2),$$
$$n_B^{acc'''}(R)=n_B^{alc''}(R)$$
$$\approx 4\left(\frac{z_{\frac{\alpha}{2}}}{l}\right)^2(\sqrt{v_1(\boldsymbol{\eta})}\sqrt{\frac{a_A}{a_A+b_A}}+\sqrt{v_2(\boldsymbol{\eta})}\sqrt{\frac{a_B}{a_B+b_B}})\times$$
$$\sqrt{v_2(\boldsymbol{\eta})}\sqrt{\frac{a_B+b_B}{a_B}}-(a_B+b_B-2)$$

The interpretation of Result 4.5.3 is limited, as it provides an upper bound on the expected number of people developing the disease and not the sample size. Nevertheless, its significance remains the same as before, in that it can serve as an approximation of the sample size or a grid search around it can yield the precise sample size with reduced computational and time complexity.

One important thing to be noted is that the ESS for prior (4.3.3), computed using the methods described in Section 2.1.1, is close to $a_A+b_A+a_B+b_B$. Interestingly, it is seen from Results 4.5.1, 4.5.2, and 4.5.3, that the total sample sizes are roughly equal to some quantity *minus* $a_A+b_A+a_B+b_B-4$, which is roughly equal to the ESS. For example, we can see from Result 4.5.2 that $n^{acc'''}(N)$ is the difference of $\left(\frac{z_{\frac{\alpha}{2}}}{l}\right)^2(\sqrt{v_1(\boldsymbol{\eta})}+\sqrt{v_2(\boldsymbol{\eta})})^2-(a_A+b_A+a_B+b_B-4)$ and ESS. The first component may be looked upon as the total amount of information needed to achieve a precision l when the coverage probability is α. This, when subtracted from the ESS, which measures the amount of information available from the trial, gives the amount of additional information needed to successfully conduct the trial at the given level of precision. This is nothing but the total required sample size, or the number of individuals to be recruited so that the additional information can be gathered. Therefore, it can be loosely stated that the Bayesian approach *bridges the gap* between the ESS and the sample size determination problem.

4.5.4 Some Simulation Results

It is quite evident that apart from length l and coverage probability α, the Bayesian sample sizes and the optimal allocation ratios are primarily functions of the hyper-parameters. In this section, we obtain the sample sizes for various values of the hyper-parameters under various allocation techniques (Equal, Double, Neyman and RSIHR). In Bayesian sample size determination, the hyper-parameters can be chosen in such a way that they deal with the uncertainties associated the underlying assumed values of VE and ARU without altering them. It should be noted that the frequentist sample sizes are functions of the desired precision l, coverage probability α and the assumed values of VE and ARU. Suppose the investigator wants to ensure that the maximum width of the 95% CI/HPD interval is 0.4 on an average, i.e. we have $l = 0.4$ and $\alpha = 0.05$. We have some prior information that the anticipated VE and ARU values are 0.4 and 0.1, respectively. Now, there is some uncertainty associated with this information in the sense that it may or may not come from a reliable source, and the authenticity of the information may be questionable. The frequentist approach has no way to deal with this uncertainty; one simply plugs in the values of VE and ARU in equation (3.2.6) to obtain the required sample sizes under Equal and Double allocation as 1704 and 1746. But, if the information is wrong and the assumed values vary widely from the actual values, we may end up with completely disastrous results that lead to a very high sample size or low precision. The Bayesian approach can counter this flaw by choosing the hyper-parameters a_A, b_A, a_B and b_B in such a way that the uncertainty is reflected in the amount of information contained in the prior. The hyper-parameters are chosen in such a way that prior expectations are equal to the assumed values of ARU and ARV, i.e.

$$ARU = E(p_A) = \frac{a_A}{a_A + b_A} \implies b_A = a_A \frac{1 - ARU}{ARU} \tag{4.5.6}$$

$$\text{and, } ARV = (1 - VE)ARU = E(p_B) = \frac{a_B}{a_B + b_B}$$

$$\implies b_B = a_B \frac{1 - ARV}{ARV}. \tag{4.5.7}$$

Therefore, by choosing various combinations of (a_A, b_A, a_B, b_B) that satisfy equations (4.5.6), (4.5.7) and conditions of Lemma 4.5.1, we can construct infinitely many priors for p_A and p_B, all of which have the same prior mean but their variances or ESS are different. Setting the prior mean equal to VE and ARU ensure that our assumption regarding their values do not change on an average. It has already been mentioned in Chapter 2, the higher the value of ESS, the more informative the prior is. The prior ESS and variance deals with the uncertainty, and if the prior information is somewhat unreliable, we can choose a vague prior with low ESS or high variance. On the other hand, if the prior information is trustworthy, we can select a slightly informative prior

with a higher ESS. While a wrong information can lead to misleading results in the frequentist approach, the Bayesian approach can still make use of this information to get the appropriate sample size without affecting precision. Tables 4.1,4.2, 4.3 and 4.4, respectively, report the total sample size required under Equal, Double, Neyman and RSIHR allocation for priors constructed by taking various values of the hyper-parameters a_A, b_A, a_B and b_B. All the priors have the same mean, and they ensure that the assumed values of VE and ARU are equal to 0.4 and 0.1, on an average. The values of l and α are respectively 0.4 and 0.05. The 5th, 6th and 7th columns, respectively, report the ACC, ACC'' and ALC sample sizes, which have been obtained following an asymptotic approach for HPD interval computation. The ACC and ALC sample sizes using the Monte Carlo approach have been computed only for the fixed allocation techniques and shown in the 8th and 9th column of Tables 4.1 and 4.2. The last column of all the tables reports the approximate upper bounds obtained in Results 4.5.1, 4.5.2 and 4.5.3. The required sample sizes in the vaccine group have also been reported within parentheses, and the corresponding frequentist sample sizes have been reported in the last row for comparison. Note that formula (3.2.6) has been used to compute the frequentist sample sizes instead of formula (3.2.5). This is because the former has been derived by directly using the asymptotic distribution of $\widehat{\psi}$ (instead of $\widehat{\beta}$) and is more comparable to the Bayesian sample sizes that have been constructed using the HPD intervals for ψ (instead of β).

				Asymptotic			Monte Carlo		
a_A	b_A	a_B	b_B	ACC	ACC''	ALC	ACC	ALC	Approximate bound
13	117	8	125.33	2220 (1110)	2046 (1023)	1702 (851)	2192 (1096)	1714 (857)	2216 (1108)
15	135	9	141.00	2042 (1021)	1886 (943)	1612 (806)	2042 (1021)	1624 (812)	2048 (1024)
17	153	10	156.67	1900 (950)	1762 (881)	1536 (768)	1882 (941)	1548 (774)	1920 (960)
19	171	11	172.33	1784 (892)	1660 (830)	1470 (735)	1770 (885)	1476 (738)	1814 (907)
21	189	12	188.00	1684 (842)	1572 (786)	1410 (705)	1670 (835)	1422 (711)	1724 (862)
23	207	13	203.67	1596 (798)	1496 (748)	1356 (678)	1588 (794)	1366 (683)	1646 (823)
Frequentist Sample Size						1705 (852)			

Table 4.1: Table showing the required total sample sizes (both ACC and ALC) under Equal allocation using both asymptotic and Monte-Carlo approach of HPD interval computation. All the hyper-parameters have same prior mean which ensures that the assumed values of VE and ARU are respectively 0.4 and 0.1. The corresponding sample sizes in the vaccine group has been reported within parentheses, and the equivalent frequentist sample sizes has been reported in the last row.

				Asymptotic			Monte Carlo		
a_A	b_A	a_B	b_B	ACC	ACC''	ALC	ACC	ALC	Approximate bound
13	117	8	125.33	2364 (1576)	2154 (1436)	1737 (1158)	2352 (1568)	1770 (1180)	2394 (1596)
15	135	9	141.00	2145 (1430)	1959 (1306)	1632 (1088)	2133 (1422)	1665 (1110)	2184 (1456)
17	153	10	156.67	1974 (1316)	1809 (1206)	1545 (1030)	1965 (1310)	1575 (1050)	2025 (1350)
19	171	11	172.33	1833 (1222)	1689 (1126)	1470 (980)	1830 (1220)	1494 (996)	1896 (1264)
21	189	12	188.00	1716 (1144)	1587 (1058)	1401 (934)	1716 (1144)	1428 (952)	1788 (1192)
23	207	13	203.67	1611 (1074)	1497 (998)	1338 (892)	1614 (1076)	1365 (910)	1695 (1130)
Frequentist Sample Size				1746 (1164)					

Table 4.2: Table showing the required total sample sizes under Double allocation using both asymptotic and Monte-Carlo approach of HPD interval computation. All the hyper-parameters have same prior mean which ensures that the assumed values of VE and ARU are respectively 0.4 and 0.1. The corresponding sample sizes in the vaccine group has been reported within parentheses, and the equivalent frequentist sample sizes has been reported in the last row.

a_A	b_A	a_B	b_B	ACC	ACC''	ALC	Approximate bound
13	117	8	125.33	2210 (1027)	2032 (1104.5)	1675 (950)	2211 (1158)
15	135	9	141.00	2027 (923)	1867 (1033.5)	1582 (908)	2040 (1086)
17	153	10	156.67	1882 (840)	1738 (977.5)	1504 (872)	1909 (1030)
19	171	11	172.33	1762 (773)	1633 (932)	1436 (842)	1801 (984)
21	189	12	188.00	1659 (715)	1543 (892)	1375 (814)	1709 (944)
23	207	13	203.67	1568 (664)	1464 (857)	1318 (788)	1628 (909)
Frequentist Sample Size				1674 (952)			

Table 4.3: Table showing the required total sample sizes (both ACC and ALC) under Neyman allocation using the asymptotic approach of HPD interval computation. All the hyper-parameters have same prior mean which ensures that the assumed values of VE and ARU are respectively 0.4 and 0.1. The corresponding sample sizes in the vaccine group has been reported within parentheses, and the equivalent frequentist sample sizes has been reported in the last row.

It is obvious from the results that the bounds are conservative as far as ALC (or ACC') and ACC'' (or ALC') sample sizes are concerned. But the tables show that they are sometimes lesser than the actual ACC sample sizes. However, they are still preferable as they are easy to compute and provide a compromise between the ALC and ACC sample sizes. The tables also provide evidence that all the Bayesian sample sizes (even the bounds) behave in a similar manner. They become smaller as the ESS

a_A	b_A	a_B	b_B	ACC	ACC''	ALC	Approximate bound
13	117	8	125.33	2249 (1359)	2069 (1271)	1705 (1088)	2251 (1337)
15	135	9	141.00	2064 (1270)	1901 (1189)	1613 (1046)	2079 (1256)
17	153	10	156.67	1917 (1199)	1771 (1126)	1534 (1006)	1946 (1193)
19	171	11	172.33	1796 (1142)	1664 (1073)	1465 (972)	1837 (1141)
21	189	12	188.00	1692 (1093)	1574 (1032)	1403 (942)	1744 (1097)
23	207	13	203.67	1599 (1046)	1495 (996)	1347 (917)	1662 (1058)
Frequentist Sample Size					2581 (2167)		

Table 4.4: Table showing the required total sample sizes under RSIHR allocation using the asymptotic approach of HPD interval computation. All the hyper-parameters have same prior mean which ensures that the assumed values of VE and ARU are respectively 0.4 and 0.1. The corresponding sample sizes in the vaccine group has been reported within parentheses, and the equivalent frequentist sample sizes has been reported in the last row.

becomes higher, i.e. the prior becomes more informative, and are sometimes lesser than the frequentist one when the prior is slightly informative. Also, a comparison of Table 4.3 with Table 4.1 shows that given the set of hyper-parameters, l and α are constant, the Bayesian Neyman allocation successfully reduces the sample size. The scenario can be better illustrated with Table 4.5 which reports the total sample size required under various allocation techniques for different values of VE and ARU. For each set of values of ARU and ARV, two priors have been constructed that have means equal to these assumed values but different variances. For the sake of brevity, only the bounds (approximate sample sizes) have been reported as they behave similarly to the actual sample sizes. The proportion allocated to vaccine have been reported within parentheses. From Table 4.5, we can observe that the Bayesian Neyman allocation gives considerable reduction in sample sizes over fixed allocation, especially for higher sample sizes. The Bayesian RSIHR allocation, on the other hand, assigns more individuals to the vaccine group when the vaccine is more effective and is preferred over Equal or Double allocation when one of the primary objectives in the trial is to reduce the overall chance of a person developing the disease.

4.6 Bayesian Sample Size determination in Case-Control Studies

All the techniques described above has been developed in the light of comparative cohort trials. It has been mentioned in Chapter 1 that retrospective case-control designs are also used to estimate vaccine efficacy. Model (1.4.1) is applicable in such a study and to solve the sample size determination problem, we need to find the appropriate number of controls (n_A) and cases (n_B). The RR in case-control studies is approximated by the OR,

VE	ARU	a_A	b_A	a_B	b_B	Equal Allocation	Double Allocation	Neyman Allocation	RSIHR Allocation
0.4	0.1	5	45	4	62.67	7529 (0.5)	9273 (0.67)	7367 (0.43)	7487 (0.49)
0.4	0.1	7	63	6	94.00	4597 (0.5)	5311 (0.67)	4587 (0.48)	4665 (0.54)
0.4	0.01	5	495	4	662.67	87600 (0.5)	108476 (0.67)	85471 (0.42)	86857 (0.48)
0.4	0.01	7	693	6	994.00	52303 (0.5)	60800 (0.67)	52139 (0.47)	53019 (0.54)
0.6	0.1	5	45	4	96.00	3919 (0.5)	4543 (0.67)	3908 (0.47)	4124 (0.59)
0.6	0.1	7	63	6	144.00	2418 (0.5)	2630 (0.67)	2415 (0.52)	2555 (0.64)
0.6	0.01	5	495	4	996.00	44931 (0.5)	52548 (0.67)	44735 (0.47)	47197 (0.58)
0.6	0.01	7	693	6	1494.00	27172 (0.5)	29810 (0.67)	27154 (0.51)	28717 (0.63)

Table 4.5: Table showing the required total sample sizes (approximate bounds) under various allocation techniques for different values of the hyper-parameters. The prior mean of every set of hyper-parameters guarantees that, on average, the presumed values of VE and ARU correspond to the values presented in column 1 and column 2, respectively. The proportion allocated to vaccine has been reported within parentheses.
Table generated using the approximate bounds obtained in Results 4.5.1 to 4.5.3

and parameter of interest, VE, becomes

$$1-\xi,$$

where $\xi = \frac{p_B(1-p_A)}{p_A(1-p_B)}$ is the OR.

In Chapter 3, we obtained the frequentist sample size (formulas (3.5.4) and (3.5.5)) by fixing the length of the $100(1-\alpha)\%$ CI for VE (or equivalently ξ). Evidently, the Bayesian approach considers the HPD intervals for ξ instead of its CI. Like the comparative cohort trials, use various criterion functions like ACC and ALC to obtain the various forms of the sample sizes. The method for obtaining ACC and ALC sample sizes are identical to the methods described in Sections 4.3, 4.4 and 4.5; with the only exception that the parameter of interest over here is ξ instead of ψ.

To begin with, we consider Beta priors on p_A and p_B, i.e.,

$$\begin{aligned} p_A &\sim Beta(a_A, b_A), \text{ and} \\ p_B &\sim Beta(a_B, b_B), \end{aligned}$$

independently of each other. The form the posterior density of ϕ is given by,

$$h(\phi \mid y_{An_A}, y_{Bn_B}, \boldsymbol{\eta}) = \frac{\phi^{a_B^*-1}}{B(a_A^*, b_A^*)B(a_B^*, b_B^*)} \int_0^1 \frac{x^{a_A^*+a_B^*-1}(1-x)^{b_A^*+b_B^*-1}}{(1-x+\phi x)^{a_B^*+b_B^*}} dx$$

$$= \begin{cases} \frac{B(a_A^*+a_B^*,b_A^*+b_B^*)}{B(a_A^*,b_A^*)B(a_B^*,b_B^*)}\phi^{a_B^*-1} \\ \times {}_2F_1(a_A^*+a_B^*, a_B^*+b_B^*, a_A^*+b_A^*+a_B^*+b_B^*; 1-\phi) \quad \text{if } 0<\phi<1, \\ \frac{B(b_A^*+b_B^*,a_A^*+a_B^*)}{B(a_A^*,b_A^*)B(a_B^*,b_B^*)}\phi^{-b_B^*-1} \\ \times {}_2F_1(b_A^*+b_B^*, a_B^*+b_B^*, a_A^*+b_A^*+a_B^*+b_B^*; 1-\frac{1}{\phi}) \quad \text{if } \phi \geq 1, \end{cases} \tag{4.6.1}$$

where $a_A^* = a_A + y_{An_A}$, $b_A^* = b_A + n_A - y_{An_A}$, $a_B^* = a_B + y_{Bn_B}$, $b_B^* = b_B + n_B - y_{Bn_B}$, and $\boldsymbol{\eta} = (a_A, b_A, a_B, b_B)'$ is the vector of hyper-parameters (Marshall, 1988).

Even though the exact form of the posterior density is known, it cannot be used for fulfilling the next objective which consists of obtaining the coverage of the HPD interval of length l, $\alpha_l^*(y_{An_A}, y_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta})$ and length of the $100(1-\alpha)\%$ HPD interval, $l_{1-\alpha}^*(y_{An_A}, y_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta})$. Hence, we again resort to Monte Carlo and asymptotic techniques for approximating the length and coverage of the HPD intervals so that $acc(n_A, n_B, \boldsymbol{\eta})$ and $alc(n_A, n_B, \boldsymbol{\eta})$ can be obtained. Here, $acc(n_A, n_B, \boldsymbol{\eta})$ and $alc(n_A, n_B, \boldsymbol{\eta})$ are defined along the same line as equations (4.3.5) and (4.3.6), i.e. they, respectively, represent the expected coverage of the HPD interval of length l and the expected length of the $100(1-\alpha)\%$ HPD interval, where the expectation is taken over the marginal distribution of $(y_{An_A}, y_{Bn_B})'$. The joint marginal distribution of y_{An_A} and y_{Bn_B} is

$$f(y_{An_A}, y_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta}) = \binom{n_A}{y_{An_A}} \frac{Beta(a_A + y_{An_A}, n_A + b_A - y_{An_A})}{Beta(a_A, b_A)} \times \binom{n_B}{y_{Bn_B}} \frac{Beta(a_B + y_{Bn_B}, n_B + b_B - y_{Bn_B})}{Beta(a_B, b_B)},$$

$y_{An_A} = 0, 1, \ldots, n_A$ and $y_{Bn_B} = 0, 1, 2, \ldots, n_B$. The marginal distribution takes the same form as in equation (4.3.4), but its interpretation is different over here.

4.6.1 Monte Carlo Methods for Approximating the HPD Intervals

HPD interval approximation using Monte-Carlo methods in case-control studies have been described in details in M'Lan et al. (2006). The unimodality of the posterior distribution of ξ allows for Monte-Carlo methods to be well-applied in this case.

We generate N independent observations from the posterior distribution of ϕ. Note that, if $\{p_{A1}, p_{A2}, \ldots, p_{AN}\}$ and $\{p_{A1}, p_{A2}, \ldots, p_{AN}\}$ respectively denote N iid samples from the posterior distribution of p_A and p_B for given $(y_{An_A}, y_{Bn_B})'$, then

$$\xi_i = \frac{p_{Bi}(1-p_{Ai})}{(1-p_{Bi})p_{Ai}}, \; i = 1, 2, \ldots, N,$$

represent iid samples of size N from (4.6.1). Then $\alpha_l^*(y_{An_A}, y_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta})$, $l_{1-\alpha}^*(y_{An_A}, y_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta})$, and subsequently $acc(n_A, n_B, \boldsymbol{\eta})$ and $alc(n_A, n_B, \boldsymbol{\eta})$ can be estimated as,

$$\widehat{\alpha}_l^*(y_{An_A}, y_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta}) = \max_{1\le i\le N} \frac{\#\{1 \le j \le N : \xi_{(i)} \le \xi_j \le \xi_{(i)} + l\}}{N},$$

$$\widehat{l}_{1-\alpha}^*(y_{An_A}, y_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta}) = \min_{1\le i\le N-[(1-\alpha)N]} (\xi_{(i+[(1-\alpha)N])} - \xi_{(i)}),$$

$$\widehat{acc}(n_A, n_B, \eta) = \sum_{y_{An_A}=0}^{n_A} \sum_{y_{Bn_B}=0}^{n_B} \widehat{\alpha}_l^*(y_{An_A}, y_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta}) f(y_{An_A}, y_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta}),$$

$$\widehat{alc}(n_A, n_B, \eta) = \sum_{y_{An_A}=0}^{n_A} \sum_{y_{Bn_B}=0}^{n_B} \widehat{l}_{1-\alpha}^*(y_{An_A}, y_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta}) f(y_{An_A}, y_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta}).$$

4.6.2 Asymptotic Methods for Approximating HPD Intervals

The asymptotic posterior distribution of ξ can be derived using the Bernstein Von Mises Theorem in a manner similar to that of Appendix C.1. Its form is given by

$$\phi \mid y_{An_A}, y_{Bn_B} \overset{a}{\sim} N\left(\widetilde{\phi}, w(\widetilde{p_A}, \widetilde{p_B})\right), \tag{4.6.2}$$

where $\tilde{p_A} = \frac{a_A + y_{An_A}}{a_A + b_A + n_A}$, $\tilde{p_B} = \frac{a_B + y_{Bn_B} - 1}{a_B + b_B + n_B}$, $\tilde{\phi} = \frac{\tilde{p_B}(1-\tilde{p_A})}{(1-\tilde{p_B})\tilde{p_A}}$, and $w(\tilde{p_A}, \tilde{p_B}) = \frac{\tilde{p_B}^2(1-\tilde{p_A})}{\tilde{p_A}^3(1-\tilde{p_B})^2(a_A+b_A+n_A)} + \frac{\tilde{p_B}(1-\tilde{p_A})^2}{\tilde{p_A}^2(1-\tilde{p_B})^3(a_B+b_B+n_B)}$.

Hence, for large samples the posterior distribution can well be approximated by (4.6.2), and we have,

$$\begin{aligned} acc(n_A, n_B, \boldsymbol{\eta}) &= \sum_{y_{An_A}=0}^{n_A} \sum_{y_{Bn_B}=0}^{n_B} \left\{ 2\Phi\left(\frac{l}{2\sqrt{w(\widetilde{p_A}, \widetilde{p_B})}}\right) - 1 \right\} f(y_{An_A}, y_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta}) \\ &= E_{\boldsymbol{y}} \left\{ 2\Phi\left(\frac{l}{2\sqrt{w(\widetilde{p_A}, \widetilde{p_B})}}\right) - 1 \right\}, \end{aligned} \tag{4.6.3}$$

$$\begin{aligned} alc(n_A, n_B, \boldsymbol{\eta}) &= \sum_{y_{An_A}=0}^{n_A} \sum_{y_{Bn_B}=0}^{n_B} 2z_{\frac{\alpha}{2}} \sqrt{w(\widetilde{p_A}, \widetilde{p_B})} f(y_{An_A}, y_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta}) \\ &= E_{\boldsymbol{y}} \left\{ 2z_{\frac{\alpha}{2}} \sqrt{w(\widetilde{p_A}, \widetilde{p_B})} \right\}, \end{aligned} \tag{4.6.4}$$

where $\boldsymbol{y} = (y_{An_A}, y_{Bn_B})'$. RHS of equation (4.6.3) is difficult to deal with, and to make it more manageable we often replace it by

$$acc'(n_A, n_B, \boldsymbol{\eta}) = 2\Phi\left(\frac{l}{2E_{\boldsymbol{x}}(\sqrt{w(\widetilde{p_A}, \widetilde{p_B})})}\right) - 1 \tag{4.6.5}$$

$$\text{or, } acc''(n_A, n_B, \boldsymbol{\eta}) = 2\Phi\left(\frac{l}{2\sqrt{E_{\boldsymbol{x}}(w(\widetilde{p_A}, \widetilde{p_B}))}}\right) - 1. \tag{4.6.6}$$

The resultant sample sizes, respectively, called ACC' and ACC'' sample sizes, are often used in place of the actual ACC sample sizes.

4.6.3 Determination of Sample Size

The sample size determination problem not only considers finding the total sample size at a given level of precision but also the number of controls and cases to be taken. Hence, the control-to-case ratio, denoted by c ($c = \frac{n_A}{n_B}$), is of utmost importance over here. Like the comparative cohort trials, the problem can be divided into two parts; (i) c is completely specified and (ii) when c is optimally determined keeping some relevant objective in mind.

Sample Size Determination when c is Completely Specified

If the control-to-case ratio c is a known quantity, we define the following sets in a Bayesian scenario:

$$\begin{aligned}
\mathcal{M}_1 &= \{(n_A, n_B) : acc(n_A, n_B, \boldsymbol{\eta}) \geq 1 - \alpha\}, \\
\mathcal{M}_2 &= \{(n_A, n_B) : alc(n_A, n_B, \boldsymbol{\eta}) \leq l, \\
\text{and, } \mathcal{M}' &= \{(n_A, n_B) : n_A = cn_B\},
\end{aligned}$$

Thereafter, ACC and ALC, respectively, searches for the pairs $(n_A^{acc}(f), n_B^{acc}(f)) \in \mathcal{M}_1 \cap \mathcal{M}'$ and $(n_A^{alc}(f), n_B^{alc}(f)) \in \mathcal{M}_2 \cap \mathcal{M}'$, such that

$$\begin{aligned}
(n_A^{acc}(f), n_B^{acc}(f)) &= \underset{(n_A, n_B) \in \mathcal{M}_1 \cap \mathcal{M}'}{\arg\min} \; n_A \\
(n_A^{alc}(f), n_B^{alc}(f)) &= \underset{(n_A, n_B) \in \mathcal{M}_2 \cap \mathcal{M}'}{\arg\min} \; n_A
\end{aligned}$$

As the equations (4.6.3) to (4.6.6) depend on the hyper-parameters $\boldsymbol{\eta}$, it is difficult to get their closed form solutions. Hence, we obtain certain bounds on the sample sizes (ACC', ACC'' and ALC), which act in the same manner as the original sample sizes, and can be utilized as seed values in sample size computations or as approximations of the sample sizes. Like comparative cohort trials, we begin by providing an upper bound on the expected value of the variance, $w(\widetilde{p_A}, \widetilde{p_B})$.

Lemma 4.6.1. *If the hyper-parameters a_A, b_A, a_B and b_B are such that $a_A > 3$, $b_A > 1$,*

$a_B > 1$ *and* $b_B > 3$, *then*

$$E_{\boldsymbol{x}}(w(\tilde{p_A}, \tilde{p_B})) \leq \frac{(a_A + b_A + n_A)}{(a_A + b_A + n_A - 1)(a_A + b_A + n_A - 2)}.w_1(\boldsymbol{\eta}) + \frac{(a_B + b_B + n_B)}{(a_B + b_B + n_B - 1)(a_B + b_B + n_B - 2)}.w_2(\boldsymbol{\eta}),$$

where $w_1(\boldsymbol{\eta}) = \frac{(a_A+b_A-1)(a_A+b_A-2)b_A}{(a_A-1)(a_A-2)(a_A-3)}.\frac{a_B(a_B+1)}{(b_B-1)(b_B-2)}$ *and* $w_2(\boldsymbol{\eta}) = \frac{a_A(a_A+1)}{(b_A-1)(b_A-2)}.\frac{a_B(a_B+b_B-1)(a_B+b_B-2)}{(b_B-1)(b_B-2)(b_B-3)}$.

Further, if n_A *and* n_B *are large enough such the* $n_A \longrightarrow \infty$, $n_B \longrightarrow \infty$ *and* $\frac{n_A}{n_A+n_B} \longrightarrow \frac{c}{1+c}$, *then*

$$E_{\boldsymbol{y}}(w(\tilde{p_A}, \tilde{p_B})) \lesssim w(\boldsymbol{\eta}, n_A, n_B),$$

where $w(\boldsymbol{\eta}, n_A, n_B) = \frac{w_1(\boldsymbol{\eta})}{(a_A+b_A+n_A-2)} + \frac{w_2(\boldsymbol{\eta})}{(a_B+b_B+n_B-2)}$.

Proof. The proof is similar to that of Lemma 4.5.1 given in Appendix C.2. □

We define the following quantities before stating the result for the approximate bounds. The proof of the result has been omitted as it is similar to the proof of Result 4.5.1.

$$\begin{aligned}
acc'''(n_A, n_B, \boldsymbol{\eta}) &= 2\Phi\left(\frac{l}{2\sqrt{w(\boldsymbol{\eta}, n_A, n_B)}}\right) - 1,\\
alc'(n_A, n_B, \boldsymbol{\eta}) &= 2z_{\frac{\alpha}{2}}\sqrt{E_{\boldsymbol{y}}(w(\widetilde{p_A}, \widetilde{p_B}))},\\
alc''(n_A, n_B, \boldsymbol{\eta}) &= 2z_{\frac{\alpha}{2}}\sqrt{w(\boldsymbol{\eta}, n_A, n_B)},\\
\mathcal{M}_1' &= \{(n_A, n_B) : acc'(n_A, n_B, \boldsymbol{\eta}) \geq 1 - \alpha\},\\
\mathcal{M}_2' &= \{(n_A, n_B) : alc'(n_A, n_B, \boldsymbol{\eta}) \leq l\},\\
\text{and, } (n_A^{acc'}(f), n_B^{acc'}(f)) &= \underset{(n_A,n_B)\in\mathcal{M}_1'\cap\mathcal{M}}{\arg\min}\ n_A
\end{aligned}$$

Analogously, we can also define $\mathcal{M}_1''$, $\mathcal{M}_1'''$, $\mathcal{M}_2''$ and $(n_A^{acc''}(f), n_B^{acc''}(f))$, $(n_A^{acc'''}(f), n_B^{acc'''}(f))$, $(n_A^{alc''}(f), n_B^{alc''}(f))$ based on $acc''(n_A, n_B, \boldsymbol{\eta})$, $acc'''(n_A, n_B, \boldsymbol{\eta})$, $alc''(n_A, n_B, \boldsymbol{\eta})$.

Result 4.6.1. *If the hyper-parameters satisfy the conditions of Lemma 4.6.1, then*

(i) $n_A^{acc'}(f) \leq n_A^{acc''}(f) \leq n_A^{acc'''}(f)$

(ii) $n_A^{alc}(f) \leq n_A^{alc'}(f) \leq n_A^{alc''}(f)$

(iii) Further, if the hyper-parameters are such that the quadratic equation

$$\frac{w_1(\boldsymbol{\eta})}{a_A + b_A - 2 + x} + \frac{cw_2(\boldsymbol{\eta})}{c(a_B + b_B - 2) + x} = \frac{1}{4}\left(\frac{l}{z_{\frac{\alpha}{2}}}\right)^2 \tag{4.6.7}$$

has at least one positive root, greater than or equal to 1, then

$$n_A^{acc'''}(f) = n_A^{alc''}(f) \approx n_A^{sol}(\boldsymbol{\eta}, \alpha, l, c),$$

where $n_A^{sol}(\boldsymbol{\eta}, \alpha, l, c)$ *is the maximum root of (4.6.7).*

The three popular choices of c mentioned in Chapter 3, viz. $c = 1$ (Equal Allocation), $c = 2$ (Double Allocation), and $c = 4$ (Quadruple allocation), were explored in the performed simulations. Tables 4.6, 4.7, and 4.8, respectively, report the total sample size required in a case-control study under Equal, Double and Quadruple allocation when the assumed values of VE and ARU are 0.4 and 0.1. Precision (l) and coverage probability (α) are, respectively, 0.4 and 0.05. Both the Monte-Carlo and asymptotic methods for HPD interval computation have been used, and approximate bounds have also been reported. The hyper-parameters have again been constructed in such a way that they satisfy the conditions of Lemma 4.6.1 and their means are equal to the assumed values of ARV and ARU. For comparison with the frequentist approach, we plug-in the same values of VE, ARU, l and α in formulae (3.5.5) and report them in the last row. We see that Bayesian sample sizes preserve all the nice features of the corresponding frequentist sample sizes while handling the uncertainties associated with the unknown values of p_A and p_B.

				Asymptotic			Monte Carlo		
a_A	b_A	a_B	b_B	ACC	ACC''	ALC	ACC	ALC	Approximate bound
13	117	8	120	2802 (1401)	2558 (1279)	2052 (1026)	2762 (1381)	2066 (1033)	2752 (1376)
15	135	9	135	2570 (1285)	2350 (1175)	1944 (972)	2538 (1269)	1960 (980)	2538 (1269)
17	153	10	150	2390 (1195)	2192 (1096)	1856 (928)	2364 (1182)	1872 (936)	2374 (1187)
19	171	11	165	2244 (1122)	2064 (1032)	1782 (891)	2226 (1113)	1798 (899)	2244 (1122)
21	189	12	180	2120 (1060)	1958 (979)	1714 (857)	2104 (1052)	1728 (864)	2136 (1068)
23	207	13	195	2014 (1007)	1866 (933)	1654 (827)	1996 (998)	1672 (836)	2042 (1021)
Frequentist Sample Size				1948 (974)					

Table 4.6: Table showing the required total sample sizes (both ACC and ALC) in a case-control study under Equal allocation using both asymptotic and Monte-Carlo approach of HPD interval computation. All the hyper-parameters have same prior mean which ensures that the assumed values of VE and ARU are respectively 0.4 and 0.1. The corresponding sample sizes in the vaccine group has been reported within parentheses, and the equivalent frequentist sample sizes has been reported in the last row.

				Asymptotic			Monte Carlo		
a_A	b_A	a_B	b_B	ACC	ACC''	ALC	ACC	ALC	Approximate bound
13	117	8	120	3195 (1065)	2931 (977)	2403 (801)	3144 (1048)	1549.5 (516.5)	3112.5 (1037.5)
15	135	9	135	2949 (983)	2710.5 (903.5)	3044 (1522)	3872 (1936)	1960 (980)	3852 (1926)
17	153	10	150	2755.5 (918.5)	2541 (847)	2914 (1457)	3624 (1812)	1872 (936)	3622 (1811)
19	171	11	165	2596.5 (865.5)	2403 (801)	2798 (1399)	3420 (1710)	1798 (899)	3436 (1718)
21	189	12	180	2461.5 (820.5)	2286 (762)	2696 (1348)	3244 (1622)	1728 (864)	3278 (1639)
23	207	13	195	2343 (781)	2184 (728)	2602 (1301)	3092 (1546)	1672 (836)	3140 (1570)
Frequentist Sample Size						2346 (782)			

Table 4.7: Table showing the required total sample sizes (both ACC and ALC) in a case-control study under Double allocation using both asymptotic and Monte-Carlo approach of HPD interval computation. All the hyper-parameters have same prior mean which ensures that the assumed values of VE and ARU are respectively 0.4 and 0.1. The corresponding sample sizes in the vaccine group has been reported within parentheses, and the equivalent frequentist sample sizes has been reported in the last row.

Sample Size Determination when c is Optimally Determined

Although, any relevant objective function can be optimized, we focus on finding a Bayesian version of the Neyman allocation derived in Chapter 3 (see equation (3.5.6)), which minimizes the total number of cases and controls.

To obtain the optimal c and sample size in the Bayesian case, ACC and ALC, respectively, searches for the pairs $(n_A^{acc}(N), n_B^{acc}(N)) \in \mathcal{M}_1$ and $(n_A^{alc}(N), n_B^{alc}(N)) \in \mathcal{M}_2$, such that,

$$(n_A^{acc}(N), n_B^{acc}(N)) = \underset{(n_A,n_B)\in\mathcal{M}_1}{\arg\min} \{n_A + n_B\}, \tag{4.6.8}$$

$$(n_A^{alc}(N), n_B^{alc}(N)) = \underset{(n_A,n_B)\in\mathcal{M}_2}{\arg\min} \{n_A + n_B\}. \tag{4.6.9}$$

The optimum control-to-case ratio in ACC and ALC are respectively given by

$$c_{acc}(N) = \frac{n_A^{acc}(N)}{n_B^{acc}(N)} \text{ and } c_{alc}(N) = \frac{n_A^{alc}(N)}{n_B^{alc}(N)}.$$

To obtain some bounds on the total sample size, we define the pairs $(n_A^{acc'}(N), n_B^{acc'}(N))$, $(n_A^{acc''}(N), n_B^{acc''}(N))$ and $(n_A^{acc'''}(N), n_B^{acc'''}(N))$ by replacing $\mathcal{M}_1$ in equation (4.6.8) by $\mathcal{M}_1'$,

				Asymptotic			Monte Carlo		
a_A	b_A	a_B	b_B	ACC	ACC''	ALC	ACC	ALC	Approximate bound
13	117	8	120	4426 (885)	4068 (814)	3384 (677)	4345 (869)	3380 (676)	4280 (856)
15	135	9	135	4095 (819)	3774 (755)	3215 (643)	4030 (806)	3215 (643)	3984 (797)
17	153	10	150	3831 (766)	3543 (709)	3073 (615)	3771 (754)	3070 (614)	3751 (750)
19	171	11	165	3611 (722)	3350 (670)	2948 (590)	3560 (712)	2945 (589)	3559 (712)
21	189	12	180	3421 (684)	3185 (637)	2834 (567)	3375 (675)	2835 (567)	3394 (679)
23	207	13	195	3255 (651)	3040 (608)	2729 (546)	3210 (642)	2733 (547)	3248 (650)
Frequentist Sample Size						3430 (686)			

Table 4.8: Table showing the required total sample sizes (both ACC and ALC) in a case-control study under Quadruple allocation using both asymptotic and Monte-Carlo approach of HPD interval computation. All the hyper-parameters have same prior mean which ensures that the assumed values of VE and ARU are respectively 0.4 and 0.1. The corresponding sample sizes in the vaccine group has been reported within parentheses, and the equivalent frequentist sample sizes has been reported in the last row.

$\mathcal{M}_1''$ and $\mathcal{M}_1'''$ respectively. Similarly, $(n_A^{alc'}(N), n_B^{alc'}(N))$ can be defined by replacing $\mathcal{M}_2$ in equation (4.6.9) by $\mathcal{M}_2'$. The optimal allocation ratio, viz. $c_{acc'}(N)$, $c_{acc''}(N)$, $c_{acc'''}(N)$ and $c_{alc'}(N)$, can be also defined corresponding to each of the above pairs.

Result 4.6.2. *(i) Under the conditions of Lemma 4.6.1 on the hyper-parameters,*

$$n_A^{acc'}(N) + n_B^{acc'}(N) \leq n_A^{acc''}(N) + n_B^{acc''}(N) \leq n_A^{acc'''}(N) + n_B^{acc'''}(N).$$

(ii) $n_A^{alc}(N) + n_B^{alc}(N) \leq n_A^{alc'}(N) + n_B^{alc'}(N) n_A^{alc''}(N) + n_B^{alc''}(N)$

(iii)
$$\begin{aligned} n_A^{acc'''}(N) + n_B^{acc'''}(N) &= n_A^{alc''}(N) + n_B^{alc''}(N) \\ &\approx 4\left(\frac{z_{\frac{\alpha}{2}}}{l}\right)^2 (\sqrt{w_1(\boldsymbol{\eta})} + \sqrt{w_2(\boldsymbol{\eta})})^2 - (a_A + b_A + a_B + b_B - 4) \\ \rho_{acc'''}(N) &= \rho_{alc'}(N) \\ &\approx \frac{4\left(\frac{z_{\frac{\alpha}{2}}}{l}\right)^2 \sqrt{w_1(\boldsymbol{\eta})}(\sqrt{w_1(\boldsymbol{\eta})} + \sqrt{w_2(\boldsymbol{\eta})}) - (a_A + b_A - 2)}{4\left(\frac{z_{\frac{\alpha}{2}}}{l}\right)^2 (\sqrt{w_1(\boldsymbol{\eta})} + \sqrt{w_2(\boldsymbol{\eta})})^2 - (a_A + b_A + a_B + b_B - 4)} \end{aligned}$$

Proof. Proof is similar to the proof of Result 4.5.2 □

The sample sizes required under Neyman allocation in a case-control study with the same setup as in Tables 4.6, 4.7 and 4.8 have been reported in Table 4.9. Only the asymptotic methods have been used for HPD interval computation, as applying Monte-Carlo methods here are too computationally extensive. The bounds have also been provided, and we see that the Neyman allocation provides the least sample size as compared to Equal, Double and Quadruple allocation. However, just like its frequentist counterpart, the Bayesian Neyman allocation also suffers from the drawback of occasionally selecting more cases than controls, which is undesirable.

a_A	b_A	a_B	b_B	ACC	ACC''	ALC	Approximate bound
13	117	8	120	2797 (1451)	2553 (1337)	2036 (1112)	2752 (1408)
15	135	9	135	2564 (1348)	2341 (1246)	1926 (1062)	2534 (1316)
17	153	10	150	2381 (1268)	2180 (1176)	1836 (1021)	2369 (1245)
19	171	11	165	2232 (1202)	2050 (1119)	1759 (987)	2237 (1188)
21	189	12	180	2106 (1146)	1941 (1071)	1691 (956)	2126 (1140)
23	207	13	195	1996 (1097)	1846 (1029)	1628 (928)	2030 (1098)
Frequentist Sample Size					1926 (1066)		

Table 4.9: Table showing the required total sample sizes under Neyman allocation using the asymptotic and Monte-Carlo approach of HPD interval computation. All the hyper-parameters have same prior mean which ensures that the assumed values of VE and ARU are respectively 0.4 and 0.1. The corresponding sample sizes in the vaccine group has been reported within parentheses, and the equivalent frequentist sample sizes has been reported in the last row.

4.7 Real-Life Example

In this section, we compare our sample sizes to the actual sample sizes taken in a trial. Here, we assume that the values of VE, ARU and the desired precision (l) are taken from the estimates of a real-life trial. In order to facilitate comparison with the frequentist approach, the same dataset as in Chapter 3, which considered the pooled interim findings from four blinded, randomized controlled trials reported in Voysey et al. (2021), was taken. Recall that the vaccine efficacy was calculated for five instances, viz., all LD/SD and SD/SD recipients, SD/SD recipients, LD/SD recipients, any symptomatic COVID-19 disease, and any NAAT-positive swab (see Section 3.6 for details). For all the five situations, the data related to case occurrence and the actual sample sizes taken in the trial were used to get the estimates of VE, ARU, ARV and even the 95% CI for VE. In Chapter 3, we calculated the required sample sizes under various allocation schemes by considering the estimated values of VE and ARU to be their underlying values, and the obtained width (l) of the CI for VE were considered to be the desired level of precision. However, to calculate the Bayesian sample sizes we need to fix the set of hyper-parameters ($\boldsymbol{\eta}$) in addition to the quantities VE, ARU and l. The hyper-parameters are set in such

a way that the prior mean of ARV and ARU coincide with their assumed values. We consider two sets of priors that have the same mean but different variances so that one is on the non-informative side and the other on the informative one. The value of the hyper-parameters along with the values of VE, ARU and precision (l) have been reported in the first and second columns of Table 4.10. For each kind of situation, we have four rows which reports the required sample sizes under the different allocation schemes. The first row reports the required sample sizes under the lesser informative prior; while the second reports the sample sizes for the prior which contains more information. The corresponding frequentist sample sizes (using formulae (3.5.5)) for the same values of VE, ARU and l, has been reported in the third row for the comparison purpose. Note that, the numbers reported here may slightly differ from those mentioned in Table 3.7, as the former used formula (3.5.4) for computation while the latter considers formula (3.5.5). The last row finally reports the actual sample sizes that were taken in the trial.

Table 4.10 shows that when the prior is more informative, the required sample sizes are sometimes lesser than the actual sample size. The less informative prior, however, safeguards against any misinformation regarding the values of VE and ARU by taking larger sample sizes. The purpose of this section was to compare the Bayesian sample sizes to the actual sample sizes keeping the estimates of VE and ARU unchanged. However, only the estimates of VE and ARU have been used in setting the prior means but no information from the trial has been used in setting the prior variances. However, in reality, the standard error estimates of ARU and ARV give an idea about the uncertainties associated with their estimates and may be used in setting the prior variance of a future trial. For example, the estimates from a Phase II trial may serve as prior information for the upcoming Phase III trial.

4.8 Concluding Remarks

In this chapter, we have calculated the HPD intervals for VE using asymptotic and Monte-Carlo methods. The Monte Carlo method is favored only when the sample sizes are small, as it presents computational challenges. The asymptotic approach, on the other hand, is in line with the frequentist approach which uses the asymptotic distribution of VE (or ψ) to obtain the CI for VE.

As this chapter is intended to provide a Bayesian alternative to the problem of determining sample sizes as discussed in Chapter 3, one of its primary objectives was to assess whether the Bayesian approach could offer any additional benefit over the traditional frequentist approach. The advantage of the Bayesian approach in terms of handling the uncertainties associated with the assumed values of the unknown parameters is similar to any other sample size determination problem and has already been discussed earlier.

	Hyper-parameters (a_A, b_A, a_B, b_B)	Equal Allocation	Double Allocation	Neyman Allocation	RSIHR Allocation
All LD/SD and SD/SD recipients $VE = 0.702$ $ARU = 0.017$ $l = 0.242$	(13, 737.27, 14, 2695.93)	12081 (6041)	12438 (8292)	11946 (6681)	13329 (9728)
	(15, 850.69, 16, 3081.07)	10743 (5371)	11112 (7408)	10650 (5893)	11955 (8768)
	Frequentist Sample Size	11624 (5812)	10701 (7134)	10685 (6926)	11605 (8952)
	Actual Sample Size	11636 (5807)			
SD/SD recipients $VE = 0.618$ $ARU = 0.016$ $l = 0.337$	(11.5, 710.08, 12.5, 2043.06)	10184 (5092)	10853 (7235)	10158 (5366)	10907 (7323)
	(13.5, 833.58, 14.5, 2369.94)	8846 (4423)	9432 (6288)	8831 (4626)	9525 (6442)
	Frequentist Sample Size	8888 (4444)	8494 (5663)	8409 (5208)	8878 (6435)
	Actual Sample Size	8895 (4440)			
LD/SD recipients $VE = 0.899$ $ARU = 0.022$ $l = 0.238$	(5, 224, 5, 2273.33)	3605 (1802)	4014 (2676)	3594 (1673)	5641 (4670)
	(6, 268.8, 6, 2728)	1975 (987)	2340 (1560)	1845 (546)	3408 (2835)
	Frequentist Sample Size	2735 (1368)	2235 (1490)	2149 (1636)	2724 (2478)
	Actual Sample Size	2741 (1367)			
Any symptomatic COVID-19 disease $VE = 0.668$ $ARU = 0.019$ $l = 0.245$	(14.5, 740.15, 16.5, 2573.11)	11760 (5880)	12201 (8134)	11658 (6439)	12780 (9086)
	(16.5, 842.24, 18.5, 2885)	10691 (5346)	11128 (7419)	10617 (5814)	11689 (8342)
	Frequentist Sample Size	11625 (5812)	10869 (7246)	10823 (6884)	11606 (8730)
	Actual Sample Size	11636 (5807)			
Any NAAT-positive swab $VE = 0.554$ $ARU = 0.026$ $l = 0.253$	(16, 593.57, 18, 1519.15)	13040 (6520)	13688 (9125)	12948 (7060)	13559 (8900)
	(17, 630.67, 19, 1603.54)	12631 (6316)	13245 (8830)	12541 (6843)	13140 (8646)
	Frequentist Sample Size	11627 (5814)	11383 (7589)	11169 (6716)	11611 (8047)
	Actual Sample Size	11636 (5807)			

Table 4.10: Table showing the total (vaccine) sample sizes required (computed using both Bayesian and frequentist approaches) under the different allocation schemes. The values of VE and ARU are taken from the estimates of the real-life trial, and the precision l is the length of the CI for VE obtained from the trial. The actual sample sizes in the trial have been reported for comparison.

The added gain was that the Bayesian method was equally successful as its frequentist counterpart if one deflects from the conventional equal (or fixed) allocation technique to other optimal allocation schemes. It was observed that the acquired sample sizes exhibit comparable behavior to the equivalent frequentist sample sizes. They increase with the increase in the underlying value of VE or decrease in ARU, provided the other factors

viz. precision, coverage probability etc are kept constant (see Table 4.5). Again, when the vaccine is more efficacious or the disease attack rate in the general population is low, the Bayesian Neyman allocation showcase substantial reduction in sample size. On the other hand, RSIHR allocation fulfill the ethical considerations in a trial by assigning more individuals to the vaccine group when the vaccine is highly effective.

One disadvantage of the Bayesian method is that the sample sizes are significantly influenced by the prior hyper-parameters; even minor adjustments to the prior have an impact on both the sample size and the optimal allocation ratio. This limitation can be addressed with the help of the bounds, which can be easily calculated and effectively reflect the changes in sample size. But, even though the bounds can be used to get an idea about sample size, it can be seen from the results that the prior hyper-parameters must satisfy certain conditions for the bounds to exist. It might be the case that the selected prior based on the available information do not satisfy the required conditions. In such cases, one either has to make adjustments in the prior hyper-parameters, or go through an intensive computational procedure to get the Bayesian sample sizes. Hence, great care must be taken during prior elicitation, as it plays a pivotal role in Bayesian sample size determination.

Chapter 5

Sample Sizes for Time-to-Event Vaccine Trial Data

5.1 Introduction

Chapters 3 and 4 have dealt thoroughly with the problem of sample size determination in vaccine trials. However, the basic model assumption in both the chapters remains the same (model (1.3.1)). The number of cases of the disease in the vaccine and placebo groups are assumed to be binomially distributed with parameters p_A and p_B. In Chapter 1, we have already pointed out the flaw in the basic definition of VE (equation (1.2.1)), and that the binomial model may not suffice in providing an idea about the true effectiveness of the vaccine. The estimates of p_A (or ARU) and p_B (or ARV), obtained using the binomial model, are actually the estimates of $ARU(L)$ and $ARV(L)$, taken at the end of the study. Keeping in mind the dynamic nature of the vaccine trials, it was suggested that a survival model, which truly identifies the time-to-event nature of the data, would be statistically more appropriate.

In a survival framework, that definition of vaccine efficacy is modified in terms of hazard ratio and denoted by $VE_H(t)$ (see equation (1.6.2)). We have also discussed the advantage of $VE_H(t)$ over the traditionally used $VE(t)$, and how the latter may give out a false sense of waning when the true $VE_H(t)$ is time-independent. In this chapter, we assume $VE_H(t)$ to be independent of time, i.e., for some constant γ, $\gamma(t) = \gamma$ for all t (see equation (1.6.3) for the expression of $\gamma(t)$). This assumption is valid because even after following the binomial model, historically there has been evidence of vaccine trials that did not show significant signs of waning, at least during the study period. For example, the ZOE-50, a randomized placebo-controlled trial, was conducted to evaluate the efficacy of a subunit vaccine containing varicella-zoster virus glycoprotein E and the $AS01_B$ adjuvant system (called HZ/su) in reducing the risk of herpes zoster in older adults (aged $\geq$ 50

years). Participants across 18 countries were enrolled, and each participant was followed for 3 to 4 years (Lal et al., 2015). Throughout the study, the HZ/su vaccine maintained a consistent efficacy of around 97%, demonstrating minimal to no signs of waning. Since, like most trials, ZOE-50 also calculated vaccine efficacy in terms of the incidence rates adjusted per 1000 person-years of follow-up instead of hazard rates, one may wonder whether this assumption actually holds in practice. However, we have already discussed in Chapter 1 how $VE(t)$ may introduce a false sense of waning, and it is expected that a vaccine showing minimal waning in $VE(t)$ will perform similarly, if not better, in terms of $VE_H(t)$. Therefore, we can safely consider the assumption of time independence of $VE_H(t)$, as it is likely to hold in certain (but not all) real-life trials.

It is quite obvious that changing the underlying model assumption to a survival one would affect the sample size calculations done before the beginning of the trial. The Neyman and RSIHR allocation proportions calculated assuming binomial model will change even though the motive behind their development will not. In this Chapter, we tackle the sample size determination problem by considering a survival model on the time-to-event data of the vaccine trials. A comparison of the results with the traditional binomial sample sizes was done, and it clearly showed the superiority of the survival approach. Here, we begin by finding a $100(1-\alpha)\%$ CI of $VE_H(=1-e^{\gamma})$, the common value of $VE_H(t)$ under the time-independent assumption, using the asymptotic distribution of some estimate of γ. In a spirit similar to that of the binomial model, the sample size determination problem in survival models is approached by fixing the width (or relative width) of the CI for VE_H and subsequently finding the smallest size required to achieve that width. One can then proceed to derive allocations analogous to Neyman and RSIHR.

The rest of the Chapter is organized as follows. In Section 5.2, we introduce the survival model and state various assumptions considered on the distributions of the survival and censoring times. A semi-parametric approach that applies the Cox Proportional Hazards (PH) model and does not require any distributional assumptions on the survival time has also been outlined. In Section 5.3, we determine the sample sizes in the survival set up by considering various choices of the survival and censoring time distributions. The sample sizes are calculated by fixing the width of the confidence interval (CI) of the time-invariant VE_H. In Section 5.4, we obtain the sample sizes by fixing the power of the relevant test for VE_H instead of the width of its CI. In Section 5.5, we take the help of a real-life trial and compare our results to those considered in the trial. In Section 5.6, we present the methods for achieving the derived allocation ratios in a trial. In Section 5.7, we briefly discuss the changes that would occur if we consider a mixture distribution to the time to censoring. Some concluding remarks are mentioned in Section 5.8.

5.2 The Survival Model

Time-to-event datasets are affected by the presence of censoring, and one should consider distributions on both censoring and survival times while modeling them.

5.2.1 Censoring Distribution

We consider only right-censored observations in our calculations as it is the most common form of censoring encountered in vaccine trials due to loss of follow-up, drop-out from the study, or study termination. The censoring time of an individual in the study is assumed to be independent of its survival time and unaffected by its reception of a vaccine or placebo. Further, it is presumed that each person is followed up to a fixed period of time, say L, provided the observation is not censored before due to some other reason. Under ideal conditions, when all the individuals are recruited at once at the beginning of the study, L simply becomes equal to the duration of the study. If C denotes the common censoring time of an individual in either of the two groups, it is quite plausible to assume that $C \leq L$ with probability 1.

A distribution of C which is degenerate at L (i.e., $C \sim \delta(L)$) is the ideal scenario where censoring is absent due to all external reasons other than the fact that there is some natural censoring, as some subjects may experience the event after the follow-up period. Uniform distribution, $U[0, L]$, are often considered on C to explain the underlying censoring mechanism (Zhang & Rosenberger, 2007). In reality, censoring may occur due to both the above phenomenon, and a perfect censoring distribution should be a mixture of the two. In other words,

$$C \sim \theta U[0, L] + (1 - \theta)\delta(L), \tag{5.2.1}$$

where θ is the mixing weight ($0 \leq \theta \leq 1$). When there are a few drop-outs or losses to follow-up and censoring predominantly occurs due to study termination, a small θ is preferred. However, its value may be adjusted based on whichever component is more dominant. We consider the two extreme cases viz. $\theta = 0$ and $\theta = 1$, and they respectively refer to the Natural and Uniform censoring mechanism. The sample sizes in the general case can be derived in a similar manner, and it is expected that they will somewhat lie between the two extremes. A natural intuition would be to incorporate θ as a parameter in the model and estimate it. Section 5.7 provides a short discussion on this model.

5.2.2 Distribution of the Survival Times

Suppose T_A and T_B respectively denote the survival time of an individual in the placebo and vaccine groups. Note that, $\lambda_A(t)$ and $\lambda_B(t)$ are the respective hazard functions of T_A and T_B. Let,

$$X_J = \min(T_J, C) \text{ and } \Delta_J = \begin{cases} 1 & \text{if } X_J \leq C \\ 0 & \text{otherwise} \end{cases}, J = A, B.$$

We have n_J observations on (X_J, Δ_J), viz. $\{(X_{J1}, \Delta_{J1}), (X_{J2}, \Delta_{J2}), \ldots, (X_{Jn_J}, \Delta_{Jn_J}), J = A, B\}$. Hence, a total of n $(n = n_A + n_B)$ individuals are recruited, of which a proportion $\rho = \frac{n_A}{n}$ is allocated to placebo. Exponential and Weibull are the two commonly used distributions in survival analysis which often fits properly or approximately to real-life data. Here, T_A and T_B are assumed to be independent, and we consider two kinds of distributions on them, viz.

1. $T_J \sim Exp(\lambda_J)$, the exponential distribution with rate λ_J, and
2. $T_J \sim Weibull(\lambda_J, p)$, the Weibull distribution with scale parameter λ_J and shape parameter p, $J = A, B$.

In both the above cases, the hazard ratios $(\frac{\lambda_B(t)}{\lambda_A(t)})$ are independent of t, and we can get closed form expressions for the asymptotic distribution of the estimate of γ under both natural and uniform censoring.

5.2.3 The Proportional Hazards Model

It is interesting to note that one can estimate γ even without making any distributional assumptions on T_A and T_B. This is due to the fact that there is proportional hazards interpretation to (1.6.3) as the equation

$$\lambda_B(t) = \lambda_A(t)e^{\gamma} \tag{5.2.2}$$

holds true under time-independent vaccine efficacy. In other words, assuming VE_H to be time-independent is equivalent to making the proportional hazards (PH) assumption. It is easy to see that equation (5.2.2) coincides with the Cox PH model (Cox, 1972) when there is a single binary covariate in the data. The single covariate, denoted by Z, in our case is the vaccination indicator. This means that $Z = 1$ or 0 according to whether one receives the vaccine or not.

In fact, we could assume Exponential and Weibull distributions to survival times in Section 5.2.2 as their families are closed under proportionality of hazards. It is due to this proportinality of hazards interpretation that any other distribution can be assumed on T_A and T_B as long as they satisfy the PH assumption.

However, the primary advantage of the PH framework is that one can estimate γ using the Cox PH model which is a semi-parametric model and does not require the parametric form of the baseline hazard ($\lambda_A(t)$ in our case) to be specified. Here, γ can be estimated by maximizing the partial likelihood, and the asymptotic distribution of the estimate can also be obtained (Cox, 1975; Tsiatis, 1981; Andersen & Gill, 1982).

5.3 Determination of Sample Size in Survival Data

In Section 5.2, we observed that γ can be estimated either by presuming distributions on the survival times (fully parametric approach) or by using the Cox PH model (semi-parametric approach). The parametric approach can further be subdivided into two parts based on the consideration of Exponential or Weibull distributions on T_A and T_B. For all three cases, we obtain the minimum sample size required to estimate γ under both Natural and Uniform censoring.

5.3.1 Exponential Survival Times

If T_J's are exponentially distributed, the likelihood of the data is given by,

$$L(\lambda_A, \lambda_B) \propto \prod_{J=A,B} \prod_{i=1}^{n_J} \{f_J(x_{Ji})^{\Delta_{Ji}} S_J(x_{Ji})^{1-\Delta_{Ji}}\}, \tag{5.3.1}$$

where $f_J(x_{Ji}) = \lambda_J e^{-\lambda_J x_{Ji}}$, $\lambda_J > 0$ is the probability density function (pdf) of T_J, and $S_J(x_{Ji}) = e^{-\lambda_J x_{Ji}}$ is the survival function, $J = A, B$.

It is well-known that, here

$$\begin{pmatrix} \widehat{\lambda}_A^{exp} \\ \widehat{\lambda}_B^{exp} \end{pmatrix} \overset{a}{\sim} N_2 \left(\begin{pmatrix} \lambda_A \\ \lambda_B \end{pmatrix}, \mathbf{I}(\lambda_A, \lambda_B)^{-1} \right),$$

where $\widehat{\lambda}_J^{exp} = \frac{\sum_{i=1}^{n_J} \Delta_{Ji}}{\sum_{i=1}^{n_J} x_{Ji}}$ is the maximum likelihood estimate (MLE) of λ_J, and

$$\mathbf{I}(\lambda_A, \lambda_B) = \begin{bmatrix} \frac{n_A e_A}{\lambda_A^2} & 0 \\ 0 & \frac{n_B e_B}{\lambda_B^2} \end{bmatrix}$$

is the Fisher information matrix. Note that,

$$\begin{aligned} e_J &= P(\Delta_J = 1) \\ &= \begin{cases} 1 - e^{-\lambda_J L} & \text{if } C \sim \delta(L) \text{ (Natural censoring)} \\ 1 - \frac{1}{\lambda_J L}(1 - e^{-\lambda_J L}) & \text{if } C \sim U[0, L] \text{ (Uniform censoring)} \end{cases}, \; J = A, B, \end{aligned} \tag{5.3.2}$$

is the probability that an individual in the respective vaccine or placebo group develops the disease within the follow-up period.

The asymptotic distribution of the MLE of γ, $\widehat{\gamma}_{exp} = \log \frac{\widehat{\lambda}_B^{exp}}{\widehat{\lambda}_A^{exp}}$ and subsequently, the $100(1-\alpha)\%$ CI of $VE_H (= 1 - e^{\gamma})$ can, respectively, be derived as,

$$\begin{aligned} \widehat{\gamma}_{exp} &\overset{a}{\sim} N(\gamma, \sigma^2_{exp}), \\ \text{and } (1 - e^{\widehat{\beta}+\widehat{d}}, 1 - e^{\widehat{\beta}-\widehat{d}}), \end{aligned}$$

where $\sigma^2_{exp} = \frac{1}{e_A n_A} + \frac{1}{e_B n_B}$ is the asymptotic variance of $\widehat{\gamma}_{exp}$ under exponential model, and $d = z_{\frac{\alpha}{2}} \sigma_{exp}$ is the asymptotic half-width of the CI of γ. The asymptotic width (W) and relative width (RW) of the CI for VE_H can then be obtained as,

$$W(\gamma, d) = (1 - VE_H)(e^d - e^{-d}) \tag{5.3.3}$$

$$\text{and, } RW(\gamma, d) = \frac{(1 - VE_H)}{VE_H}(e^d - e^{-d}). \tag{5.3.4}$$

Equations (5.3.3) and (5.3.4) are similar to the W and RW (see equations (3.2.3) and (3.2.4)) obtained in Chapter 3, but the variance terms and their interpretations are different. However, similar to the binomial case, specifying the maximum value of W/RW beforehand is equivalent to fixing the maximum half-width, d, of the CI for γ. For a given d, the minimum total sample size required to achieve this precision is given by

$$n_{exp} = \left(\frac{z_{\frac{\alpha}{2}}}{d}\right)^2 \left\{\frac{1}{\rho e_A} + \frac{1}{(1-\rho)e_B}\right\}. \tag{5.3.5}$$

Equation (5.3.5) may be looked upon as a function of ρ, and we call it $h(\rho)$. As in the case of binomial model, we simply put in the known value of ρ in $h(\rho)$ to get the required sample size under fixed allocation (like Equal or Double allocation) techniques.

Like Section (3.3.1), for a prefixed constant $k(> 0)$, we consider the optimization problem

$$\min n_A + n_B \text{ subject to}$$

$$\frac{1}{e_A n_A} + \frac{1}{e_B n_B} = k \tag{5.3.6}$$

to obtain an allocation analogous to Neyman allocation. Solving (5.3.6), we get the optimum ρ that minimizes the total sample size as

$$\boxed{\rho^{exp}_{Neyman} = \frac{\sqrt{e_B}}{\sqrt{e_B} + \sqrt{e_A}}.} \tag{5.3.7}$$

Alternatively, obtaining the ρ that minimizes $h(\rho)$ yields the same allocation proportion as (5.3.7).

The RSIHR-like allocations developed in literature after assuming exponential survival times usually minimizes the total expected hazard, $n_A\lambda_A + n_B\lambda_B$ (Zhang & Rosenberger, 2007; Biswas et al., 2016; Das & Bhattacharya, 2022). Even though such an allocation can be developed here as well, it will be of little relevance later on when we consider Weibull distributions on survival times or follow the semi-parametric approach. This is because, unlike the exponential case, the hazard functions $\lambda_A(t)$ and $\lambda_B(t)$ are usually time-dependent. Moreover, the RSIHR allocation was developed in vaccine trials with the objective of reducing the expected number of people developing the disease without compromising the precision of the study. An analogous allocation in our case would aim to minimize the expected number of events occurring within the study period, given by $n_A e_A + n_B e_B$. In case of Natural censoring, it basically minimizes the number of events occurring before time L, and has direct reference to the allocation proposed by Biswas & Mandal (2004) that reduces the expected number of responses occurring before a threshold. Therefore, similar to Section (3.3.2), the optimal allocation proportion, obtained by minimizing

$$\begin{aligned} &n_A e_A + n_B e_B \text{ subject to} \\ &\frac{1}{e_A n_A} + \frac{1}{e_B n_B} = k \end{aligned} \tag{5.3.8}$$

is given by

$$\boxed{\rho^{exp}_{RSIHR} = \frac{e_B}{e_B + e_A}.} \tag{5.3.9}$$

Again, if we minimize

$$n_A e_A + n_B e_B = h(\rho)\rho e_A + h(\rho)(1-\rho)e_B$$

with respect to ρ, we get the same ρ^{exp}_{RSIHR} as the solution.

Overall, we see that the sample sizes come out to be a function of the desired W

or RW of the CI, the coverage probability α, and the probabilities of event occurrences within the study period, e_A and e_B. Here, e_A and e_B are, in turn, functions of the hazard rates λ_A and λ_B (or equivalently λ_A and VE_H), and the length of the follow-up period L.

Tables 5.1 and 5.2 report the total sample size required to achieve a RW of 0.5 and 1 on the 95% CI when the underlying values of VE_H are assumed to be 0.4 and 0.8 respectively. The hazard rate in the unvaccinated population λ_A have been taken to be 0.01, 0.02 and 0.05, whereas the values of L are taken to be 10, 20 and 30. These values have respectively been reported in the first and second column. Columns 3 to 6 report the required sample sizes for the four allocation techniques: Equal, Double, Neyman and RSIHR, presuming Natural censoring while columns 7 to 10 reports the same under Uniform censoring. The allocation proportions ρ for the fixed allocation approaches, Equal and Double, are $\frac{1}{2}$ and $\frac{1}{3}$, respectively, whilst its value for Neyman and RSIHR allocation have been calculated using formulas (5.3.7) and (5.3.9). The corresponding sample sizes in the vaccine group have been reported within parentheses.

It can be seen from the tables that the sample sizes decrease drastically with the decrease in λ_A or increase in VE_H, provided all other factors are kept constant. These observations are in line with the binomial case where the sample sizes decrease with the decrease in ARU (disease attack rate in the unvaccinated population) or increase in VE. However, two other factors influence the survival sample size: the duration of the follow-up period (L) and the level of censoring in the data. The sample sizes are comparatively smaller when the follow-up time is smaller or the amount of censoring is bare minimal (natural censoring). These phenomena can intuitively be explained by the fact that when the study period is small or the chance of getting censored is high, there is a high possibility of loss of information on each individual. Therefore, in order to achieve the same level of precision, additional subjects must be recruited so that the loss can be compensated.

Tables 5.1 and 5.2 even pinpoint the advantages of Neyman and RSIHR allocation over the conventional fixed allocation techniques. At a given level of precision, the Neyman allocation provides fewer samples than the Equal or Double allocation, and the gain in terms of reduction of sample size is more prominent for higher sample sizes. Incidentally, the total sample sizes under Equal and RSIHR allocations are nearly the same, but it was seen that the sample size in the vaccine group is higher under RSIHR allocation for larger values of VE_H. This occurs as the RSIHR allocation ensures that a lesser number of people experience the event during the study on average. Although similar observations were made in the binomial approach as well, the advantage of the survival approach over the binomial one is that the latter may severely underestimate or overestimate the sample sizes (see Table 5.3).

As mentioned in Chapter 1, the binomial model apparently calculates vaccine efficacy

					$RW = 1$				
		Natural Censoring				Uniform Censoring			
λ_A	L	Equal	Double	Neyman	RSIHR	Equal	Double	Neyman	RSIHR
0.01	10	1983	2052	1954	1983	3918	4049	3859	3918
		(992)	(1368)	(1096)	(1230)	(1959)	(2699)	(2168)	(2436)
	20	1029	1068	1015	1029	2008	2079	1979	2008
		(514)	(712)	(567)	(634)	(1004)	(1386)	(1109)	(1242)
	30	711	741	702	711	1371	1423	1353	1371
		(356)	(494)	(391)	(435)	(686)	(949)	(756)	(845)
0.02	10	1029	1068	1015	1029	2008	2079	1979	2008
		(514)	(712)	(567)	(634)	(1004)	(1386)	(1109)	(1242)
	20	553	578	547	553	1053	1096	1040	1053
		(277)	(385)	(303)	(336)	(527)	(730)	(579)	(646)
	30	396	416	392	396	736	769	728	736
		(198)	(277)	(216)	(237)	(368)	(513)	(403)	(447)
0.05	10	459	481	454	459	863	899	852	863
		(229)	(320)	(250)	(276)	(431)	(600)	(474)	(527)
	20	272	289	270	272	484	509	479	484
		(136)	(193)	(146)	(159)	(242)	(339)	(263)	(289)
	30	213	229	212	213	359	381	356	359
		(106)	(153)	(113)	(121)	(179)	(254)	(194)	(210)
					$RW = 0.5$				
0.01	10	7726	7995	7613	7726	15262	15774	15031	15262
		(3863)	(5330)	(4271)	(4793)	(7631)	(10516)	(8446)	(9491)
	20	4008	4161	3954	4008	7821	8101	7708	7821
		(2004)	(2774)	(2209)	(2468)	(3910)	(5401)	(4319)	(4840)
	30	2771	2886	2737	2771	5342	5545	5269	5342
		(1386)	(1924)	(1523)	(1694)	(2671)	(3697)	(2944)	(3290)
0.02	10	4008	4161	3954	4008	7821	8101	7708	7821
		(2004)	(2774)	(2209)	(2468)	(3910)	(5401)	(4319)	(4840)
	20	2155	2251	2130	2155	4104	4269	4050	4104
		(1077)	(1501)	(1180)	(1308)	(2052)	(2846)	(2257)	(2516)
	30	1542	1620	1527	1542	2868	2995	2834	2868
		(771)	(1080)	(840)	(923)	(1434)	(1997)	(1571)	(1742)
0.05	10	1786	1872	1767	1786	3362	3504	3320	3362
		(893)	(1248)	(976)	(1077)	(1681)	(2336)	(1845)	(2052)
	20	1060	1126	1053	1060	1884	1982	1866	1884
		(530)	(751)	(571)	(619)	(942)	(1321)	(1025)	(1125)
	30	830	892	826	830	1398	1483	1388	1398
		(415)	(594)	(441)	(470)	(699)	(989)	(754)	(819)

Table 5.1: Table showing the total sample size required to achieve a relative width (RW) of 1 and 0.5 on the 95% CI for VE_H after assuming Exponential survival times under both Natural and Uniform censoring. The equivalent sample sizes in the vaccine group have been reported within parentheses. Four allocation types viz. Equal, Double, Neyman and RSIHR have been considered and the value of VE_H has been assumed to be 0.4 throughout. The corresponding values of λ_A and L have been reported in the first two columns.

Table generated using formulae (5.3.5)

					$RW = 1$				
		Natural Censoring				Uniform Censoring			
λ_A	L	Equal	Double	Neyman	RSIHR	Equal	Double	Neyman	RSIHR
	10	225	198	197	225	447	393	392	447
0.01		(112)	(132)	(136)	(186)	(224)	(262)	(270)	(371)
	20	114	101	101	114	226	199	199	226
		(57)	(67)	(69)	(94)	(113)	(133)	(136)	(187)
	30	78 (39)	69 (46)	69 (47)	78 (63)	152	135	135	152
						(76)	(90)	(92)	(125)
	10	114	101	101	114	226	199	199	226
0.02		(57)	(67)	(69)	(94)	(113)	(133)	(136)	(187)
	20	59 (30)	53 (35)	53 (36)	59 (48)	116	102	102	116
						(58)	(68)	(70)	(95)
	30	41 (20)	37 (24)	37 (24)	41 (33)	79 (39)	70 (47)	70 (47)	79 (64)
	10	48 (24)	43 (29)	43 (29)	48 (39)	94 (47)	83 (55)	83 (56)	94 (76)
0.05	20	26 (13)	24 (16)	24 (16)	26 (20)	49 (25)	45 (30)	45 (30)	49 (39)
	30	19 (9)	18 (12)	18 (11)	19 (14)	35 (17)	32 (21)	32 (21)	35 (27)
					$RW = 0.5$				
	10	603	530	530	603	1200	1053	1051	1200
0.01		(302)	(354)	(364)	(499)	(600)	(702)	(723)	(996)
	20	307	271	271	307	607	534	533	607
		(153)	(181)	(185)	(252)	(303)	(356)	(366)	(501)
	30	208	185	184	208	409	361	361	409
		(104)	(123)	(125)	(170)	(204)	(241)	(246)	(336)
	10	307	271	271	307	607	534	533	607
0.02		(153)	(181)	(185)	(252)	(303)	(356)	(366)	(501)
	20	159	141	141	159	310	275	275	310
		(79)	(94)	(95)	(129)	(155)	(183)	(187)	(254)
	30	109	98 (66)	98 (66)	109	211	188	188	211
		(55)			(87)	(106)	(126)	(127)	(171)
	10	129	116	116	129	251	223	223	251
0.05		(65)	(77)	(78)	(104)	(125)	(149)	(151)	(204)
	20	70 (35)	64 (43)	64 (42)	70 (55)	132	120	120	132
						(66)	(80)	(79)	(106)
	30	51 (25)	48 (32)	47 (30)	51 (38)	93 (47)	85 (57)	85 (56)	93 (73)

Table 5.2: Table showing the total sample size required to achieve a relative width (RW) of 1 and 0.5 on the 95% CI for VE_H after assuming Exponential survival times under both Natural and Uniform censoring. The equivalent sample sizes in the vaccine group have been reported within parentheses. Four allocation types viz. Equal, Double, Neyman and RSIHR have been considered and the value of VE_H has been assumed to be 0.8 throughout. The corresponding values of λ_A and L have been reported in the first two columns.

Table generated using formulae (5.3.5)

using formula (1.2.1), which is nothing but $VE(t)$ (defined by equation (1.5.1) calculated at $t = L$. Here, $ARU(t)$ and $ARV(t)$ are, respectively, estimated by the proportion of

individuals developing the disease in the placebo and vaccine groups. If the observations are actually coming from a survival distribution, then a misspecified binomial model basically estimates of probabilities of the occurrence of an event before time t in the two groups. Hence we can write,

$$ARU(t) = Pr(X_A \leq t, \Delta_A = 1),$$
$$\text{and } ARV(t) = Pr(X_B \leq t, \Delta_A = 1),\ t \leq L.$$

From the above expressions, it is clear that $ARU(L)$ and $ARV(L)$ respectively reduce to e_A and e_B, the formulas of which for the exponential model with natural and uniform censoring are given by (5.3.2).

In order to compare our sample sizes to the binomial ones, we tried to compute the sample sizes that would have been obtained if one considered the binomial model instead of the actual survival model. For various values of L ($L = 8, 10, 12$), the $ARU(L)$ and $ARV(L)$ values were computed and substituted in the binomial sample size formula (equation (3.2.5)) as p_A (ARU) and p_B (ARV). The results obtained for different L's, and under various allocation techniques have been shown in Table 5.3. Column 5 of Table 5.3 basically reports the sample sizes that would have been obtained if one mistakenly used the binomial model during their calculations when the exponential model with natural censoring is the true underlying model. Column 8 basically reports the same with the only exception that it considers uniform censoring instead of natural. The corresponding sample sizes obtained by considering the true model (formula (5.3.5)) have been shown in columns 4 and 7 for comparison. The desired width of the 95% CI is fixed at 0.25, and it is kept the same in both the survival and binomial sample size calculations. The value of λ_A in the correct model and the actual vaccine efficacy VE_H are chosen to be 0.05 and 0.6, respectively. The $VE(L)$ values have also been reported which shows that vaccine efficacy will likely to be underestimated under the binomial model, especially when L is large. Table 5.3 shows that the binomial model takes fewer samples under Natural censoring, which eventually leads to a loss in precision. Conversely, the sample sizes are seen to be overestimated under Uniform censoring, which may incur certain unnecessary costs in the study. Overall, it is clear that the Binomial sample sizes provide fallacious estimates of the sample sizes.

5.3.2 Weibull Survival Times

In the next part of the parametric approach, we assume that the T_J's follow Weibull distribution with pdf

$$f_J(t) = p\lambda_J t^{p-1} e^{-\lambda_J t^p},\ \lambda_J > 0,\ p > 0,\ J = A, B.$$

		Natural Censoring			Uniform Censoring		
L	Allocation Type	$VE(L)$	Survival Sample size	Binomial Sample size	$VE(L)$	Survival Sample size	Binomial Sample size
8	Equal	0.55	795 (398)	791 (395)	0.57	1531 (766)	1588 (794)
8	Double	0.55	781 (521)	748 (498)	0.57	1495 (997)	1522 (1015)
8	Neyman	0.55	765 (458)	743 (466)	0.57	1469 (886)	1506 (929)
8	RSIHR	0.55	795 (549)	771 (551)	0.57	1531 (1070)	1569 (1115)
10	Equal	0.54	654 (327)	648 (324)	0.56	1248 (624)	1304 (652)
10	Double	0.54	645 (430)	609 (406)	0.56	1222 (814)	1248 (832)
10	Neyman	0.54	631 (376)	606 (382)	0.56	1199 (721)	1235 (763)
10	RSIHR	0.54	654 (448)	627 (449)	0.56	1248 (867)	1285 (912)
12	Equal	0.53	560 (280)	552 (276)	0.55	1059 (530)	1115 (557)
12	Double	0.53	555 (370)	516 (344)	0.55	1040 (693)	1065 (710)
12	Neyman	0.53	542 (321)	514 (326)	0.55	1019 (611)	1055 (653)
12	RSIHR	0.53	560 (380)	531 (381)	0.55	1059 (732)	1096 (776)

Table 5.3: Table showing the comparison between the sample sizes calculated using the Survival model and the mis-specified Binomial model under various allocation techniques and different values of L. The desired width of the 95% CI is kept fixed at 0.25. For the true survival model, the value of λ_A and VE_H are chosen to be 0.05 and 0.6 respectively.

The log-likelihood l up to some irrelevant constant is given by,

$$l(\lambda_A, \lambda_B, p) = \sum_{J=A,B} \log \lambda_J \sum_{i=1}^{n_J} \Delta_{Ji} + \log p \sum_{J=A,B} \sum_{i=1}^{n_J} \Delta_{Ji} + (p-1) \sum_{J=A,B} \sum_{i=1}^{n_J} \Delta_{Ji} \log x_{Ji} - \sum_{J=A,B} \lambda_J \sum_{i=1}^{n_J} x_{Ji}^p.$$

Standard calculations show that the Fisher Information matrix is given by

$$I(\lambda_A, \lambda_B, p) = \begin{bmatrix} \frac{n_A e_A}{\lambda_A^2} & 0 & n_A e_{1A} \\ 0 & \frac{n_B e_B}{\lambda_B^2} & n_B e_{1B} \\ n_A e_{1A} & n_B e_{1B} & n_A E_{2A} + n_B E_{2B} \end{bmatrix},$$

where $E_{2J} = \frac{e_J}{p^2} + \lambda_J e_{2J}$, $e_J = P(\Delta_J = 1)$, $e_{1J} = E(X_J^p \log X)$, and $e_{2J} = E(X_J^p (\log X)^2)$.

The expressions for e_J, e_{1J} and e_{2J} under both Natural and Uniform censoring have been derived and mentioned in Appendix D.1. The expressions involve computation of integrals of the form $\int_0^{L^p} z^a (\log z)^b e^{-\lambda z} dz$, and they have been obtained numerically using the `integrate` function in R. If $(\widehat{\lambda_A^{wei}}, \widehat{\lambda_B^{wei}}, \widehat{p})'$ denote the maximum likelihood estimate

(MLE) of $(\lambda_A, \lambda_B, p)$, and n_A, n_B are sufficiently large with $\frac{n_A}{n_A+n_B} \longrightarrow \rho$, then

$$\sqrt{n}(\widehat{\gamma}_{wei} - \gamma) \xrightarrow{d} N(0, \sigma^2_{wei}),$$

where $\widehat{\gamma}_{wei} = \log \frac{\widehat{\lambda}_B^{wei}}{\widehat{\lambda}_A^{wei}}$ is the MLE of γ, and

$$\sigma^2_{wei} = \frac{\{\rho E_{2A} + (1-\rho)E_{2B}\}\{\rho e_A + (1-\rho)e_B\} - \{\rho\lambda_A e_{1A} + (1-\rho)\lambda_B e_{2B}\}^2}{\rho(1-\rho)e_A e_B\{\rho E_{2A} + (1-\rho)E_{2B}\} - \rho(1-\rho)^2 e_A\lambda_B^2 e_{1B}^2 - \rho^2(1-\rho)\lambda_A^2 e_{1A}^2 e_B},$$

is the asymptotic variance.

Then, similar to Section 5.3.1, we find the W or RW of the CI for VE_H, from the asymptotic distribution of $\widehat{\gamma}_{wei}$. As before, specifying W/RW is equivalent to specifying d as they are connected through equation (5.3.3)/(5.3.4). The form of the half-width of the CI for γ, denoted by d, is different over here and is given by,

$$d = z_{\frac{\alpha}{2}} \frac{\sigma_{wei}}{\sqrt{n}}.$$

Hence, for specified d, the required sample size in the Weibull model is given by,

$$\boxed{n_{wei} = \left(\frac{z_{\frac{\alpha}{2}}}{d}\right)^2 \sigma^2_{wei}.} \tag{5.3.10}$$

The RHS is a function of ρ among other quantities, and we denote it by $H(\rho)$. Obtaining the sample size under fixed allocation techniques is quite straight forward, as the known value of ρ is simply fed into equation (5.3.10). But, to obtain Neyman and RSIHR allocation proportions, the optimization problems, (5.3.6) and (5.3.8) in this case, respectively, become

$$\begin{aligned} &\text{minimize } n_A + n_B \\ &\text{subject to } \frac{\sigma^2_{wei}}{n} = k, \end{aligned} \tag{5.3.11}$$

$$\begin{aligned} &\text{and minimize } n_A e_A + n_B e_B \\ &\text{subject to } \frac{\sigma^2_{wei}}{n} = k \end{aligned} \tag{5.3.12}$$

As the form of σ^2_{wei} is too complicated, solving (5.3.11) and (5.3.12) would be mathematically cumbersome. For simplicity, we consider the alternative interpretation that minimizes the total sample size $(n_A + n_B)$ and average number of people developing the

disease within the study period ($n_A e_A + n_B e_B$) with respect to ρ. In other words, we find

$$\rho^{wei}_{Neyman} = \arg \min_{\rho \in (0,1)} H(\rho), \tag{5.3.13}$$

$$\text{and } \rho^{wei}_{RSIHR} = \arg \min_{\rho \in (0,1)} \{\rho H(\rho) e_A + (1-\rho) H(\rho) e_B\}. \tag{5.3.14}$$

However, getting closed-form expressions for ρ^{wei}_{Neyman} and ρ^{wei}_{RSIHR} is still difficult, and we obtain them numerically using the `optimize` function in R. If $e_{1A} = e_{1B} = 0$, the forms of the Weibull and exponential models coincide and we get closed form expressions for the optimal ρ's.

The sample sizes and the optimum allocation proportions depend on the values of e_A, e_B, e_{1A}, e_{1B}, e_{2A} and e_{2B}. e_J, e_{1J}, e_{2J}, $J = A, B$, are functions of $\{\lambda_A, \lambda_B, p, L\}$ or $\{\lambda_A, VE_H, p, L\}$, and it can be said that the sample sizes are functions of the desired d, α, λ_A, VE_H, p and L. The behavior of the sample sizes was similar to that of the exponential model when the values of one of λ_A, VE_H, or L were changed, keeping the others constant. But we observed that the sample sizes initially decreased with the increase in the value of the common shape parameter p but became constant after some time. The sample sizes for two different values of p ($p = 1.5$ and $p = 2$) keeping the rest of the parameters the same, have been reported in Table 5.4. Three sets values of λ_A and L, viz. {0.01, 0.02, 0.05} and {5,10,15} have been taken, where as the specified RW and the assumed VE_H have been set at 0.4 and 0.5 respectively. We also see from the table that, like the exponential set-up, the success of Neyman and RSIHR allocations in achieving their respective objectives is more pronounced for larger sample sizes.

5.3.3 The Semi-Parametric Approach

The sample sizes obtained in Section 5.3.1 or 5.3.2 work perfectly when the underlying survival times actually follow an Exponential or a Weibull distribution, which may not always be the case. Moreover, the sample size calculations are done before the beginning of the study, and it is difficult to guess the distribution of the upcoming observations at that time. This drew us towards the semi-parametric approach, which does not require any distributional assumptions on T_A and T_B other than the fact that their hazards should satisfy the PH assumption.

In order to estimate γ using the Cox PH model, we slightly redesign the parametric framework. Hence, instead of considering n_A and n_B observations in the placebo and vaccine groups, we associate a covariate Z to all n observations. Here, Z is the vaccine indicator, which is actually a binary random variable with $P(Z = 1) = 1 - \rho$.

		$p = 1.5$							
		Natural Censoring				Uniform Censoring			
λ_A	L	Equal	Double	Neyman	RSIHR	Equal	Double	Neyman	RSIHR
0.01	5	5420	5445	5276	5420	13361	13400	12997	13361
		(2710)	(3630)	(3073)	(3580)	(6681)	(8933)	(7586)	(8856)
	10	2049	2076	2002	2049	4926	4965	4803	4926
		(1025)	(1384)	(1154)	(1331)	(2463)	(3310)	(2786)	(3231)
	15	1215	1243	1192	1215	2827	2868	2764	2827
		(608)	(829)	(679)	(773)	(1414)	(1912)	(1591)	(1831)
0.02	5	2812	2837	2743	2812	6836	6875	6658	6836
		(1406)	(1892)	(1589)	(1840)	(3418)	(4583)	(3873)	(4505)
	10	1135	1163	1114	1135	2624	2664	2566	2624
		(567)	(775)	(633)	(719)	(1312)	(1776)	(1475)	(1695)
	15	728	760	720	728	1581	1624	1554	1581
		(364)	(507)	(399)	(444)	(791)	(1082)	(880)	(998)
0.05	5	1254	1282	1230	1254	2925	2966	2860	2926
		(627)	(855)	(701)	(799)	(1463)	(1977)	(1647)	(1897)
	10	606	640	601	606	1255	1299	1237	1255
		(303)	(427)	(328)	(359)	(627)	(866)	(694)	(779)
	15	471	511	470	471	853	900	846	853
		(236)	(341)	(247)	(260)	(427)	(600)	(462)	(507)
		$p = 2$							
0.01	5	2537	2563	2476	2537	7361	7406	7171	7361
		(1268)	(1708)	(1432)	(1657)	(3681)	(4937)	(4169)	(4848)
	10	805	836	794	805	2120	2168	2080	2121
		(403)	(557)	(444)	(496)	(1060)	(1445)	(1185)	(1351)
	15	514	551	511	514	1167	1218	1153	1167
		(257)	(367)	(273)	(293)	(584)	(812)	(640)	(710)
0.02	5	1376	1403	1348	1376	3863	3909	3773	3863
		(688)	(935)	(770)	(881)	(1931)	(2606)	(2178)	(2514)
	10	540	576	537	540	1261	1311	1245	1261
		(270)	(384)	(289)	(312)	(630)	(874)	(693)	(774)
	15	432	477	432	432	803	858	799	803
		(216)	(318)	(221)	(227)	(402)	(572)	(429)	(463)
0.05	5	695	727	688	695	1774	1823	1744	1775
		(348)	(485)	(380)	(421)	(887)	(1215)	(987)	(1119)
	10	428	473	428	428	768	823	765	769
		(214)	(316)	(218)	(222)	(384)	(549)	(409)	(439)
	15	418	466	418	418	600	655	599	600
		(209)	(311)	(209)	(209)	(300)	(437)	(311)	(324)

Table 5.4: Table showing the total sample size required to achieve a relative width (RW) of 0.4 on the 95% CI for VE_H after assuming Weibull distribution with common shape parameter p on the survival times of the vaccine and placebo groups. Two values of p viz. 1.5 and 2 have been taken, and both Natural and Uniform censoring mechanisms have been considered. The corresponding sample sizes in the vaccine group under the various allocation techniques have been reported within parentheses. The value of VE_H has been assumed to be 0.5, and the values of λ_A (scale parameter of unvaccinated population) and L have been reported in the first two columns.

Table generated using formulae (5.3.10)

If T_i and C_i denote the survival time and censoring time of the ith individual, then

$$X_i = \min(T_i, C_i),\ \Delta_i = \begin{cases} 1 & \text{if } T_i \leq C_i \\ 0 & \text{otherwise} \end{cases},$$

$$\text{and, } Z_i = \begin{cases} 1 & \text{if the ith person receives the vaccine} \\ 0 & \text{otherwise.} \end{cases}$$

Hence, we observe the iid random vectors (X_i, Δ_i, Z_i), $i = 1, 2, \ldots, n$. Note that,

$$T|Z = 0 \stackrel{d}{=} T_A,\ T|Z = 1 \stackrel{d}{=} T_B,$$

and T_i has hazard $\lambda(t|z_i)$, where

$$\lambda(t|z_i) = \begin{cases} \lambda_A(t) & \text{if } z_i = 0 \text{ (baseline hazard)} \\ \lambda_B(t) & \text{if } z_i = 1. \end{cases}$$

Even though we do not assume any parametric distribution on T_i, the assumptions on the censoring distribution C_i remain the same as in Section 5.2.1. That is, we either assume C_i's, $i = 1, 2, \ldots, n$, to be iid $U[0, L]$ (Uniform censoring) or degenerate at L (Natural censoring). However, there is a basic conceptual difference between the parametric and semi-parametric framework. In the former set-up, we assumed that an exact proportion ρ of the incoming individuals are allocated to placebo, whereas the proportion is expected to be ρ over here. The slight modification is done in the latter case so that it satisfies that assumptions (Tsiatis, 1981) required to find the asymptotic distribution of the estimate of γ. However, simulation results have shown that it really doesn't matter for large n, and the asymptotic distribution remains unchanged even if we take exact proportion ρ.

The estimate of γ in the semi-parametric approach, $\widehat{\gamma}_{PH}$, is obtained by maximizing Cox's partial likelihood which takes the following form in our case:

$$PL(\beta) = \prod_{i \in \mathcal{D}} \frac{e^{\beta z_i}}{R_A(x_i) + e^{\beta} R_B(x_i)},$$

where $\mathcal{D} = \{i : \Delta_i = 1, i = 1, 2, \ldots n\}$. Also, $R_A(x_i)$ and $R_B(x_i)$, $i = 1, 2, \ldots, n$, respectively, denote the number of individuals in the placebo and vaccine groups surviving until time t.

Using the results from Tsiatis (1981), it has been shown in Appendix D.2 that

$$\sqrt{n}(\widehat{\gamma}_{PH} - \gamma) \xrightarrow{d} N(0, \sigma^2_{PH}),$$

where

$$\sigma_{PH}^{-2} = \begin{cases} \int_{S_A(L)}^{1} \frac{\rho(1-\rho)(1-VE)}{(1-\rho)(1-VE)+\rho y^{VE}} dy & \text{under Natural censoring} \\ \int_{S_A(L)}^{1} \frac{\rho(1-\rho)(1-VE)(1-\frac{S_A^{-1}(y)}{L})}{(1-\rho)(1-VE)+\rho y^{VE}} dy & \text{under Uniform censoring} \end{cases}$$

Note that, $S_A(t) = P(T_A \geq t)$ and $S_B(t) = P(T_B \geq t)$ are the survival function of the unvaccinated and vaccinated populations, respectively. Here, $\frac{\sigma_{PH}^2}{n}$ is the asymptotic variance of $\widehat{\gamma}_{PH}$, and

$$\widehat{\sigma}_{PH}^2 = \frac{1}{n} \sum_{i \in \mathcal{D}} \frac{R_A(x_i) e^{\beta} R_B(x_i)}{(R_A(x_i) + e^{\beta} R_B(x_i))^2}.$$

A $100(1-\alpha)\%$ CI for γ is, therefore, given by,

$$\widehat{\gamma} \mp \frac{\widehat{\sigma}_{PH}}{\sqrt{n}} z_{\frac{\alpha}{2}},$$

and, if we fix its asymptotic half-width to d, then the minimum sample size (total) required to achieve that width is

$$\boxed{n_{PH} = \left(\frac{z_{\frac{\alpha}{2}}}{d}\right)^2 \sigma_{PH}^2.} \tag{5.3.15}$$

Needless to say that we can alternatively obtain a $100(1-\alpha)\%$ CI for $VE_H (= 1 - e^{\gamma})$, and fix its width or relative width (with respect to VE_H) to W/RW for obtaining the sample sizes. This is possible as the one-to-one correspondence between W/RW and d applies here as well.

We again see that the required sample size (equation (5.3.15)) comes out to be a function of ρ, and in case of fixed allocation we can simply plug-in its known value to get the sample size. We can even develop optimal allocations like Neyman and RSIHR, which, respectively, minimize the total sample size n and the expected number of people developing the disease $E[\sum_{i=1}^{n} I(\Delta_i = 1)]$. Like the Weibull case, we do not have explicit expressions for the optimal allocation proportions ρ_{Neyman}^{PH} and ρ_{RSIHR}^{PH}, but we obtain them using the `optimize` function in R.

The sample sizes under both fixed and optimal allocation depend on the desired precision d, the coverage probability α, the length of the follow-up period L, and the baseline survival function $S_A(t)$. In order to make a comparison with the parametric approach, we examine two forms of the baseline hazard function: Exponential and Weibull. In precise words, the survival functions of $Exp(\lambda_A)$ and $Weibull(\lambda_A, p)$ distributions are substituted as the baseline survival function $S_A(t)$ in the sample size formula (5.3.15). Table 5.5 reports the required sample sizes under the various allocation techniques after presuming an Exponential baseline. The value of the necessary parameters RW, λ_A, L and α have been kept the same as in Table 5.1. Hence, we can say that it reports the sample sizes

one would get if they fit a Cox PH model to an Exponential dataset instead of following the regular parametric approach of Section 5.3.1. When the sample sizes are large, the ones obtained in Table 5.5 are almost similar to that of Table 5.1 showing that no additional penalty will be involved if one takes the semi-parametric approach. Similar situations were also observed when Weibull was considered to be the form of the baseline hazard. Table 5.6 reports the sample sizes obtained by following a semi-parametric approach when the underlying survival times are Weibull. The sample sizes are almost identical to those of Table 5.4, which calculates the sample sizes under the same setup by taking the parametric approach. The advantage of the semi-parametric approach is therefore clear: it works perfectly regardless of the original distribution of the survival times for higher sample sizes.

It is to be kept in mind that Tables 5.5 and 5.6 have been constructed solely for the purpose of comparison with the parametric approach. The Cox PH model was primarily developed to estimate γ when the form of the underlying baseline hazard is unknown, a phenomenon often occurring in reality. The situation is simpler in case of natural censoring where σ^2_{PH} (or n_{PH}) depend only on the value of $S_A(t)$ at $t = L$. Here, $S_A(L)$ and $S_B(L)$ are, respectively, equal to $ARU(L)$ and $ARV(L)$ under natural censoring, and guesses about their values are often the only publicly available information from previously conducted clinical trials. These information can then readily be used to get an idea about the sample size. However, in case of uniform censoring, we need information on the form of $S_A(t)$ to compute the sample sizes. In such cases, one can approximate the integral $\int_{S_A(L)}^{1} \frac{\rho(1-\rho)(1-VE)(1-\frac{S_A^{-1}(y)}{L})}{(1-\rho)(1-VE)+\rho y^{VE}} dy$ in σ^2_{PH} using the trapezoidal rule, i.e. we set

$$\int_{S_A(L)}^{1} \frac{\rho(1-\rho)(1-VE)(1-\frac{S_A^{-1}(y)}{L})}{(1-\rho)(1-VE)+\rho y^{VE}} dy \approx \frac{(1-S_A(L))}{2} \frac{\rho(1-\rho)(1-VE)}{(1-\rho)(1-VE)+\rho}$$

to get a crude estimate of the sample size. The method is not quite recommended as it may overestimate the sample sizes, but it is better than a parametric approach with completely wrong distributional assumptions. Also, if the loss-to-follow up/ drop-out rate is negligible, one need not worry about these factors, and can simply obtain the sample sizes by considering Natural censoring. Even if the form of $\lambda_A(t)$ (or $S_A(t)$) is known, a conventional parametric approach may be difficult to implement if the derivation of MLE and Fisher information matrix requires complex mathematical calculations. A semi-parametric approach is obviously preferred in such cases due to its tractable calculations.

					$RW = 1$				
		Natural Censoring				Uniform Censoring			
λ_A	L	Equal	Double	Neyman	RSIHR	Equal	Double	Neyman	RSIHR
	10	1983	2052	1954	1983	3918	4049	3859	3918
0.01		(992)	(1368)	(1097)	(1230)	(1959)	(2700)	(2168)	(2436)
	20	1029	1068	1015	1029	2008	2080	1979	2008
		(515)	(712)	(567)	(634)	(1004)	(1386)	(1109)	(1243)
	30	712	741	703	712	1372	1424	1353	1372
		(356)	(494)	(391)	(435)	(686)	(949)	(756)	(845)
	10	1029	1068	1015	1029	2008	2080	1979	2008
0.02		(515)	(712)	(567)	(634)	(1004)	(1386)	(1109)	(1243)
	20	553	578	547	553	1054	1096	1040	1054
		(277)	(385)	(303)	(336)	(527)	(731)	(580)	(646)
	30	396	416	392	396	737	769	728	737
		(198)	(278)	(216)	(237)	(368)	(513)	(404)	(448)
	10	459	481	454	459	863	900	853	863
0.05		(229)	(321)	(251)	(277)	(432)	(600)	(474)	(527)
	20	273	290	271	273	485	510	480	485
		(137)	(193)	(147)	(159)	(242)	(340)	(264)	(289)
	30	214	230	213	214	360	382	358	360
		(107)	(154)	(114)	(122)	(180)	(255)	(194)	(211)
					$RW = 0.5$				
	10	7727	7995	7613	7727	15262	15774	15032	15262
0.01		(3863)	(5330)	(4272)	(4793)	(7631)	(10516)	(8446)	(9492)
	20	4009	4162	3954	4009	7821	8101	7709	7821
		(2004)	(2774)	(2209)	(2469)	(3911)	(5401)	(4320)	(4841)
	30	2772	2887	2737	2772	5343	5546	5270	5343
		(1386)	(1925)	(1523)	(1695)	(2671)	(3697)	(2945)	(3291)
	10	4009	4162	3954	4009	7821	8101	7709	7821
0.02		(2004)	(2774)	(2209)	(2469)	(3911)	(5401)	(4320)	(4841)
	20	2156	2252	2131	2156	4105	4270	4052	4105
		(1078)	(1502)	(1181)	(1309)	(2052)	(2847)	(2258)	(2517)
	30	1544	1622	1529	1544	2870	2997	2836	2870
		(772)	(1082)	(840)	(924)	(1435)	(1998)	(1572)	(1744)
	10	1788	1873	1769	1788	3363	3506	3322	3363
0.05		(894)	(1249)	(976)	(1078)	(1682)	(2337)	(1846)	(2053)
	20	1064	1130	1056	1064	1888	1986	1870	1888
		(532)	(753)	(572)	(621)	(944)	(1324)	(1027)	(1128)
	30	835	898	832	835	1404	1489	1394	1404
		(418)	(598)	(444)	(474)	(702)	(993)	(757)	(823)

Table 5.5: Table showing the total sample size required to achieve a relative width (RW) of 1 and 0.5 on the 95% CI, following the semi-parametric approach. The corresponding sample sizes in the vaccine group have been reported within parentheses. The baseline hazard has been assumed to be Exponential with rate λ_A, and the setup along with values of the necessary parameters, viz. λ_A, L, and VE_H, have been kept the same as in Table 5.1.

Table generated using formulae (5.3.15)

					$p = 1.5$				
		Natural Censoring				Uniform Censoring			
λ_A	L	Equal	Double	Neyman	RSIHR	Equal	Double	Neyman	RSIHR
	5	5420	5445	5276	5420	13361	13400	12997	13361
0.01		(2710)	(3630)	(3073)	(3580)	(6681)	(8933)	(7586)	(8856)
	10	2050	2076	2003	2050	4926	4966	4803	4926
		(1025)	(1384)	(1155)	(1331)	(2463)	(3310)	(2786)	(3232)
	15	1216	1244	1193	1216	2828	2869	2765	2828
		(608)	(829)	(679)	(773)	(1414)	(1912)	(1591)	(1832)
	5	2812	2837	2743	2812	6836	6875	6659	6836
0.02		(1406)	(1892)	(1589)	(1840)	(3418)	(4583)	(3873)	(4505)
	10	1135	1164	1115	1135	2625	2665	2567	2625
		(568)	(776)	(633)	(719)	(1312)	(1777)	(1475)	(1696)
	15	730	761	721	730	1583	1626	1556	1583
		(365)	(507)	(400)	(445)	(792)	(1084)	(881)	(999)
	5	1254	1282	1230	1255	2926	2967	2861	2927
0.05		(627)	(855)	(701)	(799)	(1463)	(1978)	(1647)	(1897)
	10	608	642	603	609	1258	1301	1239	1258
		(304)	(428)	(329)	(361)	(629)	(868)	(695)	(781)
	15	476	516	475	476	859	905	851	859
		(238)	(344)	(250)	(264)	(429)	(603)	(466)	(510)
					$p = 2$				
	5	2537	2563	2476	2537	7362	7407	7172	7362
0.01		(1268)	(1708)	(1432)	(1657)	(3681)	(4938)	(4169)	(4848)
	10	806	837	795	807	2122	2170	2082	2123
		(403)	(558)	(444)	(497)	(1061)	(1447)	(1186)	(1352)
	15	517	554	515	517	1172	1223	1158	1172
		(259)	(369)	(275)	(295)	(586)	(815)	(642)	(714)
	5	1376	1404	1349	1376	3864	3910	3774	3864
0.02		(688)	(936)	(771)	(881)	(1932)	(2607)	(2178)	(2515)
	10	543	579	540	543	1265	1315	1249	1266
		(271)	(386)	(291)	(314)	(632)	(877)	(696)	(777)
	15	441	484	441	441	813	866	808	813
		(221)	(323)	(227)	(233)	(406)	(577)	(436)	(470)
	5	697	729	689	697	1777	1826	1746	1777
0.05		(348)	(486)	(381)	(422)	(888)	(1217)	(989)	(1121)
	10	438	481	437	438	779	832	775	779
		(219)	(320)	(225)	(230)	(389)	(554)	(416)	(447)
	15	432	476	432	432	615	666	614	615
		(216)	(317)	(221)	(221)	(308)	(444)	(322)	(336)

Table 5.6: Table showing the total sample size required to achieve a relative width (RW) of 0.4 on the 95% CI, following the semi-parametric approach. The corresponding sample sizes in the vaccine group have been reported within parentheses. The baseline hazard has been assumed to be Weibull with parameters λ_A and p, and the setup along with values of the other quantities, viz. λ_A, p, L, and VE_H, have been kept the same as in Table 5.4.

Table generated using formulae (5.3.15)

5.4 Sample Sizes Based on Fixing Power

In Section 3.4, we explored the sample size determination by fixing the power of the relevant test, instead of the precision d. There is ample literature on sample size determination in a survival framework, most of which include fixing the power of the log-rank test that detects significant difference between two survival populations (Schoenfeld, 1981; Schoenfeld & Richter, 1982; Schoenfeld, 1983; Lachin & Foulkes, 1986; Dupont & Plummer, 1990; Ahnn & Anderson, 1995; Hsieh & Lavori, 2000; Halabi & Singh, 2004). But, here we want to test for the effectiveness of the vaccine under the assumption that the underlying vaccine efficacy $VE_H(t)$ is independent of t. Hence, the pertinent hypothesis is given by,

$$H_0 : VE_H \leq 0 \text{ versus } H_1 : VE_H > 0$$

or equivalently,

$$H_0 : \gamma \geq 0 \text{ versus } H_1 : \gamma < 0.$$

It is quite obvious that the test statistic changes according to the survival and censoring time distributions considered. The assumptions on the distributions of the survival and censoring times are kept the same as in Section 5.2. Hence, the problem can be divided into two parts based on whether we take the parametric or semi-parametric approach in estimating γ. The parametric approach can further be sub-divided into two parts according to the parametric distribution (Exponential or Weibull) presumed on the survival times. We explore the three cases: Exponential, Weibull and semi-parametric, one by one and compare their results.

If one assumes the survival times of the vaccinated and unvaccinated populations to be exponentially distributed (same model as in Section 5.3.1), then the Z-test based on the large sample normal approximation of the distribution of $\widehat{\gamma}_{exp}$ reject H_0 at level α if

$$Z = \frac{\widehat{\gamma}_{exp}}{\widehat{\sigma}_{exp}} < -z_\alpha,$$

where $\widehat{\sigma}^2_{exp} = \frac{1}{\sum_{i=1}^{n_A} \Delta_{Ai}} + \frac{1}{\sum_{i=1}^{n_B} \Delta_{Bi}}$ is the estimate of σ^2_{exp}. The power function of the test is given by

$$p(\gamma) = P_\gamma(\frac{\widehat{\gamma}_{exp}}{\widehat{\sigma}_{exp}} < -z_\alpha) = 1 - \Phi(z_\alpha + \frac{\gamma}{\widehat{\sigma}_{exp}})$$

The investigator stipulates that for some γ under the alternative, the asymptotic power

of the aforementioned test must be at least $1-\eta$, i.e.

$$\begin{aligned} & 1-\Phi(z_\alpha+\frac{\gamma}{\sigma_{exp}}) \geq 1-\eta \\ \implies & n \geq \left(\frac{z_\alpha+z_\eta}{\gamma}\right)^2 \left(\frac{1}{\rho e_A}+\frac{1}{(1-\rho)e_B}\right). \end{aligned} \tag{5.4.1}$$

The quantity on the RHS of (5.4.1), denoted by n^p_{exp}, is the minimum sample size required to achieve a power (asymptotic) of $1-\eta$ under the alternative γ. The corresponding sample size when one assumes Weibull distribution on T_A and T_B with respective scale parameters λ_A and λ_B, and common shape parameter p, is given by,

$$n^p_{wei} = \left(\frac{z_\alpha+z_\eta}{\gamma}\right)^2 \sigma^2_{wei}.$$

It is quite evident that, in the Weibull scenario, one fixes the asymptotic power of the size α Z-test performed on the basis of the large sample distribution of $\widehat{\gamma}_{wei}$. The value of the alternative γ is computed from the anticipated value of VE_H, and for given γ and α, we see that both the sample sizes n^p_{exp} and n^p_{wei} are, respectively, equal to equations (5.3.5) and (5.3.10) if

$$d = \left|\frac{z_{\frac{\alpha}{2}}\gamma}{z_\alpha+z_\eta}\right|.$$

Therefore, achieving $100(1-\eta)\%$ power on the level α test is equivalent to achieving a width of $\left|\frac{z_{\frac{\alpha}{2}}\gamma}{z_\alpha+z_\eta}\right|$ on the $100(1-\alpha)\%$ CI for γ. Hence, like the standard approach, fixing power and precision are equivalent for the survival models as well, and the sample sizes behave similarly (i.e., become higher) with the increase in η or decrease in W/RW, provided the rest of the factors are kept constant. Apart from γ, the sample sizes depend on the labelling parameters of the unvaccinated population (λ_A in case of Exponential distribution; λ_A and p in case of Weibull distribution), the length of the follow-up period L, the size of the test α, the desired level of power $1-\eta$, and the allocation proportion ρ.

As the sample sizes depend on ρ, we can develop optimum allocations like Neyman and RSIHR which, respectively, minimizes $\{n_A+n_B\}$ and $\{n_Ae_A+n_Be_B\}$. Clearly, the optimum allocation proportions so obtained will take the same forms as in Section 5.3.1 (equations (5.3.7) and (5.3.9)) and Section 5.3.2 (equations (5.3.13) and (5.3.14)) based on whether we consider Exponential or Weibull survival time ditributions.

In a similar manner, if one takes the semi-parametric approach in estimating γ, H_0 is rejected at level α if

$$Z_{PH} = \sqrt{n}\frac{\widehat{\gamma}_{PH}}{\widehat{\sigma}_{PH}} < -z_\alpha.$$

The sample size needed to achieve a power of $1-\eta$ under the alternative γ is then given by

$$n_{PH}^{p} = \left(\frac{z_\alpha + z_\eta}{\gamma}\right)^2 \sigma_{PH}^2. \tag{5.4.2}$$

Like the parametric approach, there is a direct equivalence with the precision-based sample size n_{PH} in the sense that fixing η is equivalent to setting the precision d to $\left|\frac{z_{\frac{\alpha}{2}}\gamma}{z_\alpha + z_\eta}\right|$, for given γ and α. Thus, the allocation proportions under the corresponding Neyman and RSIHR allocation will remain the same as in the precision-based method (Section 5.3.3), and they can be obtained using the `optimize` function in R since their explicit forms are hard to obtain. Table 5.7 reports the required sample sizes under the different allocation techniques and for various values of the alternative VE_H. The first column of the table lists the values of the chosen alternative, while the second column displays the value of the corresponding W (obtained from equivalent precision d) for each VE_H. For example, when $VE_H = 0.3$, the value of the equivalent W is 0.24. This means that if the vaccine is 30% effective, one would get the sample size if they fix the maximum width of the 95% CI for VE_H to 0.24 instead of desiring at least 90% power in the size 0.05 test. Only Natural censoring has been considered here, because of which the sample sizes depend on the follow-up time L and the baseline survival function $S_A(t)$ only through $S_A(L)$. Two values of $S_A(L)$, viz. 0.88 and 0.91, have been considered corresponding to each VE_H, and they have been mentioned in column 3. Uniform censoring has not been considered, as it requires the exact form of the baseline hazard function $\lambda_A(t)$ to be stated. Moreover, it grossly overestimates the sample sizes and is deemed to be far from reality. For a particular VE_H, the sample sizes are always larger when $S_A(L)$ takes the greater value 0.91. This is due to the fact that under the same model, a higher $S_A(L)$ means a smaller L, and it has been mentioned earlier that the sample sizes increase with the decrease in the length of the follow-up period.

5.5 Illustration with Real-Life Data

For the purpose of comparing our results to the actual sample sizes taken in a real-life scenario, we introduce the multinational, placebo-controlled, observer-blinded, pivotal efficacy trial conducted for the BNT162b2 vaccine candidate. The trial, the findings of which are reported in Polack et al. (2020), randomly assigned participants (16 years or older) in a 1 : 1 ratio to receive two doses, 21 days apart, of either vaccine or placebo ($30\mu g$ per dose).

We aim to evaluate whether adopting Neyman or RSIHR allocation instead of the 1 : 1 allocation used in the trial would gain us a reduction in sample size or the expected

VE_H	Equivalent d	$S_A(L)$	Equal	Double	Neyman	RSIHR
0.3	0.24	0.88	2695 (1347)	2863 (1908)	2676 (1451)	2676 (1438)
		0.91	3603 (1802)	3825 (2550)	3577 (1941)	3577 (1929)
0.4	0.34	0.88	1436 (718)	1487 (992)	1415 (793)	1416 (784)
		0.91	1922 (961)	1989 (1326)	1894 (1063)	1894 (1054)
0.6	0.61	0.88	579 (290)	562 (375)	553 (337)	554 (332)
		0.91	778 (389)	753 (502)	742 (452)	742 (447)
0.7	0.81	0.88	413 (206)	383 (256)	382 (245)	382 (241)
		0.91	555 (277)	515 (343)	513 (330)	513 (326)

Table 5.7: Table showing the total sample sizes (sample sizes in the vaccine group) required to achieve a power of 90% on the size 0.05 test for VE_H under various values of the alternatives. The semi-parametric approach was followed and only the Natural censoring mechanism was considered. The corresponding precision d and $S_A(L)$ values have been stated in 2nd and 3rd columns.
Table generated using formulae (5.4.2)

number of people developing the disease, provided the unknown parameter estimates and precision of the study remain unaltered. In Section 3.6, we did a similar analysis by considering the conventional Binomial model. We obtained the estimates of VE, ARU, and the width (or relative width) of the 95% CI for VE using the trial data reported in Voysey *et al.* (2021). The estimated VE and ARU are assumed to be their true values, and the obtained value of d is taken to be the desired level of precision. These values are then plugged into formula (3.2.5) to get the required sample sizes under the various allocation techniques, and the results are subsequently compared with the actual sample size taken in the trial (see Table 3.7).

In a similar spirit, here we set the width (or relative width) of the 95% CI of VE_H, calculated from the trial data, to be the desired width. The estimates of VE_H and other relevant parameters, obtained following the parametric or semi-parametric approach, are then passed onto the relevant sample size formula(s). However, it is to be noted that to get the estimate of VE_H, one needs the complete dataset in the survival analysis framework, which includes the information on the disease development time or censoring time for every individual. Such data are publicly unavailable, and Polack et al. (2020) only reports the number of cases of the disease developed within the study period. This unavailability of data creates major hindrances in the way of an appropriate data analysis, and we try to proceed with the limited information in hand.

In the BNT162b2 vaccine trial, 43, 548 participants underwent randomization, of whom 43, 448 received injections: 21, 720 with the vaccine and 21, 728 with placebo. The first primary end point was the efficacy of BNT162b2 against confirmed COVID-19 with onset at least 7 days after the second dose in participants who has been without serologic or virologic evidence of SARS-CoV-2 infection up to 7 days after the second dose; the second primary endpoint was efficacy in participants with and without evidence of prior

infection. Among the 36,523 participants (18,198 in the vaccine group and 18,325 in the placebo group) who had no evidence of existing or prior SARS-CoV-2 infection, 8 cases of COVID-19 were observed among vaccine recipients and 162 among placebo recipients; thereby reporting a VE of 95%. On the other hand, among the 40,137 participants (19,965 in the vaccine group and 20,172 in the placebo group) with and those without prior evidence of SARS-CoV-2, 9 cases of COVID-19 at least 7 days after the second dose were observed among vaccine recipients and 169 among placebo recipients. This case split corresponds to 94.6% VE.

In both the instances described above, the efficacies were calculated following the usual Binomial model. The surveillance time was calculated in this regard, which is basically the total time in 1000 person-years for the given endpoint across all participants within each group at risk for the endpoint. For participants without prior evidence of infection, the surveillance time was 2.214 in the vaccine group and 2.222 for the placebo group. In clinical language, this means that the amount of information obtained by continuously observing individuals in the trial is equivalent to the amount of data that would have been collected if 2214 and 2222 persons in the respective vaccine and placebo groups were followed for one year. The Binomial model uses this interpretation to adjust for factors like different accrual times and the presence of censoring in the data. The disease attack rates are then adjusted in terms of per 1000 person-years of follow-up, and in our case the estimated ARV and ARU (or rather $ARV(L)$ and $ARU(L)$) come out to be $\frac{8}{2214}$ and $\frac{162}{2222}$, respectively. For the rest of the section, we will work with this adjusted data, which hypothetically assumes that 2214 and 2222 people in the vaccine and placebo groups were followed for a period of 1 year (i.e., $L = 1$). The number of cases of the disease or the number of people experiencing the event within the one-year period in the two groups are 8 and 162, and these are the only information available about the observed data. We assume Natural censoring as all other censoring information has been suppressed during the adjustment. Note that one of our primary purposes was to avoid these kinds of adjustments in the disease attack rates and fit the more statistically appropriate survival model to the data. However, the lack of information on the survival and censoring times compelled us to work with the adjusted data. Had the entire dataset been available, we could have done a better analysis using the appropriate estimate(s) of VE_H obtained from the real-life trial. Here, the estimate has been obtained in a slightly different manner, using the estimates of e_A (or $ARU(L)$) and e_B (or $ARV(L)$). We initially assume the underlying survival times to be exponentially distributed and equate e_J to its value under Natural censoring (see equation (5.3.2)) to get the estimates of λ_J, $J = A, B$, which are further used to obtain the width (W) of the 95% CI of VE_H as 0.074. We set this width as the desired width and pass it into the formula (5.3.5) to get the total sample size (under the four allocation techniques) that would have been obtained if we followed the survival approach presuming exponential survival times, provided the estimates of VE_H and λ_A

remain unchanged (see the first row of Table 5.8). Again, if we follow the semi-parametric approach, which does not require any distributional assumption on the survival times, we can find the corresponding sample sizes by simply using formula (5.3.5) instead of (5.3.15). However, the estimate of VE_H and the precision W will be obtained here differently — by plugging in the estimate of $S_J(L)(= 1 - e_J)$ in the equations

$$S_B(L) = S_A(L)^{1-VE_H} \text{ and } \sigma_{PH}^{-2} = \int_{S_A(L)}^{1} \frac{\rho(1-\rho)(1-VE_H)}{(1-\rho)(1-VE_H) + \rho y^{VE_H}} dy.$$

The resultant sample sizes along with the estimates of VE_H, $S_A(L)$, and obtained W have been reported in the second row of Table 5.8.

In order to facilitate comparison with the standard approach, the estimate and the width of the 95% CI of VE obtained by following the usual binomial model together with the total sample sizes that would have been obtained by considering these values to be the true underlying values have been reported in the third row of Table 5.8. The actual total sample size for the adjusted data, which is 4436 $(= 2214 + 2222)$, has been reported in the fourth row.

A similar analysis was also performed for the second primary efficacy endpoint, which included participants with or without evidence of prior SARS-CoV-2 infection. The disease attack rates per 1000 person-years of follow-up were reported to be 9/2332 and 169/2345, respectively. The adjusted data was obtained as before, and similar calculations were done here. The obtained results have been reported in the last four rows of Table 5.8.

We see from Table 5.8 that, like the Binomial model, the survival approaches would have also taken a lesser number of samples if Neyman allocation had been adopted instead of the actual allocation used in the trial, provided the parameter estimates and the width of the CI remain unchanged. Again, we see that since the vaccine is effective, more individuals are allocated to the vaccine group in case of RSIHR allocation. The sample sizes assuming parametric and semi-parametric models are nearly identical, whereas those presuming the Binomial model come out to be slightly lower. This shows that the Binomial model may provide wrong sample sizes even under ideal conditions. It goes without saying that the shortcomings of the Binomial model would have been more comprehensible if the complete survival (censoring) data had been available.

Fitted model	Estimated parameters	Width of the 95% CI of VE/VE_H	Equal Allocation	Double Allocation	Neyman Allocation	RSIHR Allocation
Participants without evidence of existing or prior SARS-CoV-2 infection						
Parametric survival model assuming exponential distribution on the failure times	$VE_H = 95.2\%$ $\lambda_A = 0.076$	0.074	4429 (2214)	3478 (2319)	3154 (2580)	4429 (4220)
Semi-parametric survival model only assuming proportionality of hazards	$VE_H = 95.2\%$ $S_A(L) = 0.927$	0.074	4429 (2214)	3478 (2319)	3154 (2580)	4429 (4220)
Binomial model	$VE = 95\%$ $ARU = 0.073$	0.076	4429 (2214)	3468 (2313)	3123 (2571)	4357 (4159)
Actual Sample Size			4436 (2214)			
Participants with or without evidence of prior SARS-CoV-2 infection						
Parametric survival model assuming exponential distribution on the failure times	$VE_H = 94.8\%$ $\lambda_A = 0.075$	0.075	4665 (2333)	3677 (2451)	3357 (2726)	4665 (4428)
Semi-parametric survival model only assuming proportionality of hazards	$VE_H = 94.8\%$ $S_A(L) = 0.928$	0.075	4665 (2333)	3677 (2451)	3358 (2727)	4666 (4429)
Binomial model	$VE = 94.6\%$ $ARU = 0.072$	0.077	4665 (2333)	3665 (2443)	3325 (2718)	4592 (4366)
Actual Sample Size			4667 (2332)			

Table 5.8: Table showing the sample sizes required to obtain the precision level of the study reported in Polack *et al.* (2020) if one considered the survival parametric and semi-parametric approach, provided the estimates of VE_H and other unknown parameters remain unchanged. The sample sizes derived by assuming the Binomial model, and the actual sample size (from the adjusted data) of the trial have also been mentioned for comparison. All the reported sample sizes are under the four allocation techniques, along with the corresponding sample size in the placebo group within parentheses.

5.6 Implementation

Any sample size and optimal allocation proportion determination problem can broadly be classified into two categories. The first part involves providing an initial estimate of the sample size before the beginning of the trial. However, the sample sizes often depend on the unknown parameters that are to be estimated from the trial, and investigators use the available prior information to make a guess about their values. We have already seen in the previous sections that the survival approach is no exception in this regard. The second part, in contrast, estimates the unknown parameters adaptively through the trial and ultimately works towards achieving the target allocation proportion (Neyman or RSIHR) in the long run. In Section (3.7), we observed how sequential procedures like SMLE, DBCD, and ERADE can be deployed for this purpose provided the individuals are entering the trial sequentially and the responses on all the previous patients are available at the time of enrolling a new one. The latter assumption may not be true in the case of survival outcomes, as there may be delay in responses in the sense that all the formerly entered individuals may not have experienced the event (or been censored) at the time of a new recruitment. Response Adaptive Randomization (RAR) procedures calculate the probability of allocating a new patient to the vaccine or placebo group based on the information (on allocations and responses) obtained so far, so that asymptotically a specified fraction ρ (as in (5.3.7), (5.3.9), (5.3.13), or (5.3.14)) of the recruited individuals are allocated to placebo. If there is a delay in responses, the probabilities are calculated based on the available responses. Hu et al. (2008), Bai et al. (2002) and Hu et al. (2009) showed that the theoretical results established regarding the variability of the SMLE, DBCD, and ERADE procedures remain unchanged in such a case, provided the assumption

$$Pr(\tau_m > \eta_{m+n} - \eta_m) = o(n^{-c}) \tag{5.6.1}$$

holds true for some constant $c > 0$. Here τ_m denotes the delay in response of the mth patient and η_m its survival time.

Further, Zhang & Rosenberger (2007) explored the application of RAR techniques to survival outcomes and showed that assumption (5.6.1) is satisfied when there are uniform patient entries and exponential survival times. Therefore, RAR techniques, SMLE, DBCD, and ERADE, described in Section 3.7, can be well applied in this case. However, slight adjustments will be needed in the allocation probabilities because the inherent model parameters differ in this context. For the time being, we artificially assume that there is no delay and the responses are instantly available. Participants are initially allocated to vaccine or placebo in a 1:1 manner, after which the estimates of the parameters λ_A and λ_B are obtained at every stage. Let $\widehat{\lambda}_A(j-1)$ and $\widehat{\lambda}_B(j-1)$ respectively denote the estimate of λ_A and λ_B based on the first $(j-1)$ patients; then the DBCD procedure assigns the jth patient to placebo with probability

$$g\left(\frac{N_A(j-1)}{j-1}, \rho(\widehat{\lambda}_A(j-1), \widehat{\lambda}_B(j-1)\right).$$

The definition of $N_A(j-1)$ and the form of g (see equation (3.7.1)) remain the same as before. As mentioned earlier, SMLE can be looked upon as a special case of DBCD with $\alpha_1 = 0$ in g.

On the other hand, the allocation probability of the jth individual to the placebo group takes the same form as p_j (see equation (3.7.2)) in the case of the ERADE procedure. But its meaning is slightly different, as the estimate of ρ after the jth stage, $\widehat{\rho}_{j-1}$, is $\rho(\widehat{\lambda}_A(j-1), \widehat{\lambda}_B(j-1))$ instead of $\rho(\widehat{p}_A(j-1), \widehat{p}_B(j-1))$.

The theoretical expressions for the variabilities of these designs can be obtained for exponential outcome with Natural or Uniform censoring using Theorem 1 of Zhang & Rosenberger (2007), which states that under the DBCD procedure, the asymptotic value of $nVar\left(\frac{N_A(n)}{n}\right)$ that targets an allocation proportion ρ is

$$nVar\left(\frac{N_A(n)}{n}\right) = \frac{\rho(1-\rho) + 2(1+\alpha_1)\sigma_3^2}{1+2\alpha_1}, \tag{5.6.2}$$

where $\sigma_3^2 = (\nabla\rho)'V(\nabla\rho)$, $V = diag\left\{\frac{Var(X_{A1}-\frac{\Delta_{A1}}{\lambda_A})}{\rho e_A^2}, \frac{Var(X_{B1}-\frac{\Delta_{B1}}{\lambda_B})}{(1-\rho)e_B^2}\right\}$ and $(\nabla\rho) = (\frac{d\rho}{d\lambda_A}.\lambda_A^2, \frac{d\rho}{d\lambda_B}.\lambda_B^2)'$. In case of ERADE, $nVar\left(\frac{N_A(n)}{n}\right) \longrightarrow \sigma_3^2$ as $n \longrightarrow \infty$.

Several simulation studies were performed to check whether the target allocation ratio was achieved asymptotically. We began by selecting four choices of the sample size n, viz. 150, 200, 300, and 500. For every choice of n and each of the three RAR designs, 10,000 identical trials were simulated. For example, if we take $n = 150$ and the RAR procedure to be SMLE, then in each of the simulated trials, 150 incoming subjects were randomly assigned to vaccine or placebo groups using the SMLE procedure with Neyman allocation (ρ_{Neyman}^{exp}) as the target. The survival responses were generated by assuming the underlying survival times to be exponentially distributed with rates $\lambda_A = 0.4$ and $VE_H = 0.7$ for the vaccine and placebo groups, respectively. The censoring mechanism was initially assumed to be Natural with $L = 10$, and later on the entire procedure was repeated for Uniform censoring. The simulated values of $nVar\left(\frac{N_A(n)}{n}\right)$ for different n and various RAR designs have been reported in columns 6 to 8 of Table 5.9. The corresponding simulated value of $E\left(\frac{N_A(n)}{n}\right)$, which is expected to approach the true ρ for large n, is reported within parentheses. We see that even when the sample sizes are moderate, the simulated proportions are close to ρ, with ERADE giving the best results. Theoretical values of $nVar\left(\frac{N_A(n)}{n}\right)$, computed using equation (5.6.2), are observed to be nearly equal to those of the simulated values. Similar observations were made when all the simulations were rerun by considering ρ_{RSIHR}^{exp} as the target allocation proportion (see

Table 5.10).

		Theoretical value of $nVar\left(\frac{N_A(n)}{n}\right)$			Simulated value of $nVar\left(\frac{N_A(n)}{n}\right)$ $\left(E\left(\frac{N_A(n)}{n}\right)\right)$		
n	True ρ	SMLE	DBCD	ERADE	SMLE	DBCD	ERADE
Natural Censoring							
150	0.458	0.270	0.063	0.011	0.252 (0.462)	0.061 (0.459)	0.014 (0.459)
200					0.256 (0.461)	0.062 (0.459)	0.013 (0.459)
300					0.257 (0.460)	0.064 (0.458)	0.013 (0.458)
500					0.260 (0.460)	0.064 (0.458)	0.012 (0.458)
Uniform Censoring							
150	0.427	0.310	0.088	0.033	0.267 (0.435)	0.088 (0.429)	0.035 (0.429)
200					0.282 (0.433)	0.086 (0.428)	0.034 (0.428)
300					0.288 (0.431)	0.086 (0.428)	0.034 (0.428)
500					0.300 (0.430)	0.089 (0.427)	0.034 (0.427)

Table 5.9: Table showing the theoretical and simulated variabilities of the SMLE, DBCD, and ERADE procedures with Neyman allocation proportion, ρ^{exp}_{Neyman}, as the target. The underlying model is assumed to follow exponential survival times and Natural or Uniform censoring mechanisms ($L = 10$). The values of the parameters λ_A and VE_H are chosen to be 0.4 and 0.7, respectively.

		Theoretical value of $nVar\left(\frac{N_A(n)}{n}\right)$			Simulated value of $nVar\left(\frac{N_A(n)}{n}\right)$ $\left(E\left(\frac{N_A(n)}{n}\right)\right)$		
n	True ρ	SMLE	DBCD	ERADE	SMLE	DBCD	ERADE
Natural Censoring							
150	0.416	0.322	0.096	0.039	0.281 (0.425)	0.093 (0.418)	0.042 (0.418)
200					0.293 (0.423)	0.095 (0.418)	0.040 (0.418)
300					0.301 (0.421)	0.096 (0.417)	0.041 (0.417)
500					0.309 (0.419)	0.095 (0.417)	0.039 (0.417)
Uniform Censoring							
150	0.356	0.442	0.174	0.106	0.373 (0.373)	0.165 (0.361)	0.105 (0.362)
200					0.384 (0.369)	0.171 (0.36)	0.107 (0.36)
300					0.400 (0.365)	0.171 (0.359)	0.107 (0.359)
500					0.409 (0.362)	0.177 (0.358)	0.106 (0.358)

Table 5.10: Table showing the theoretical and simulated variabilities of the SMLE, DBCD, and ERADE procedures with RSIHR allocation proportion, ρ^{exp}_{RSIHR}, as the target. The underlying model is assumed to follow exponential survival times and Natural or Uniform censoring mechanisms ($L = 10$). The values of the parameters λ_A and VE_H are chosen to be 0.4 and 0.7, respectively.

When the survival times are Weibull, the RAR procedures can be slightly modified to achieve allocation proportions ρ^{wei}_{Neyman} and ρ^{wei}_{RSIHR}. Here, the allocation proportions are functions of three parameters, viz. λ_A, λ_B, and p. The estimated ρ at any stage is calculated by replacing these parameters in the RHS equations (5.3.13) and (5.3.14) with their corresponding estimates at that stage. However, unlike the exponential distribution, there is no closed-form expression for MLE in the case of Weibull, and hence no theoretical expression of $nVar\left(\frac{N_A(n)}{n}\right)$ can be derived. Moreover, the target allocation proportions cannot be explicitly written as functions of λ_A, λ_B, and p, and hence their derivatives

with respect to the parameters cannot be obtained theoretically. But simulation studies showed that all the RAR procedures achieved the targeted allocation proportions when we generated responses with Weibull survival times. Like Figures (3.1) and (3.2), ERADE was seen to have the lowest variability for both the Exponential and Weibull distributions.

5.7 Mixture Distribution to Censoring Time

It has been discussed in Section 5.2.1 that form (5.2.1) serves as the perfect censoring distribution in real-life scenarios. However, the quantity θ $(0 \leq \theta \leq 1)$ is unknown, and treating it as a parameter would change the likelihood of the model. We have considered two extreme cases, i.e. $\theta = 0$ and $\theta = 1$, until now, but that can be easily extended to the general case, provided the value of θ is assumed to be completely specified.

If θ is known, then the form of the likelihood (5.3.1) and the information matrix in the case of the Exponential distribution are unaltered. Subsequently, the forms of σ^2_{exp}, n_{exp}, ρ^{exp}_{Neyman}, and ρ^{exp}_{RSIHR} would be the same as Section 5.3.1. Even though the forms are the same, their values will be different as e_A and e_B will be different in the case of mixture censoring distribution. Note that, for exponential survival times and censoring distribution (5.2.1),

$$e_J = \theta \left\{ 1 - \frac{1}{\lambda_J L}(1 - e^{-\lambda_J L}) \right\} + (1 - \theta)(1 - e^{-\lambda_J L}), \;\; J = A, B.$$

Evidently, the above e_J-value is a weighted mean of the e_J-values under Uniform and Natural censoring. For example, suppose we assume that the mixing weight $\theta = 0.5$, $\lambda_A = 0.01$, $VE_H = 0.4$, $L = 10$, desired relative width $RW = 1$, and $\alpha = 0.05$, then the required sample size n_{exp} under the four allocation schemes respectively becomes 2633, 2724, 2594 and 2633, respectively. As expected, the computed sample sizes fall between those obtained under the two extreme censoring mechanisms.

Similarly, in the case of Weibull distribution, the expressions for σ^2_{wei}, n_{wei}, ρ^{wei}_{Neyman}, and ρ^{wei}_{RSIHR} are identical to those in Section 5.3.2, but their values will be different as $\{e_J, e_{1J}, e_{2J}, \;\; J = A, B\}$ will vary according to the value of θ. Since $\{e_J, e_{1J}, e_{2J}, \;\; J = A, B\}$ represent estimates of various functions of the distributions of X_J, their values will clearly be the weighted mean of the corresponding values under Uniform and Natural censoring (see Appendix D.1) with weight θ.

Again, if we follow the semi-parametric approach and assume censoring distribution

(5.2.1), then from Appendix D.2, we have

$$\sigma_{PH}^{-2} = \int_{S_A(L)}^{1} \frac{\rho(1-\rho)(1-VE)S_C(S_A^{-1}(y))}{(1-\rho)(1-VE)+\rho y^{VE}}dy$$
$$= \theta \int_{S_A(L)}^{1} \frac{\rho(1-\rho)(1-VE)(1-\frac{S_A^{-1}(y)}{L})}{(1-\rho)(1-VE)+\rho y^{VE}}dy + (1-\theta)\int_{S_A(L)}^{1} \frac{\rho(1-\rho)(1-VE)}{(1-\rho)(1-VE)+\rho y^{VE}}dy$$

under the mixture censoring distribution. Note that σ_{PH}^{-2} here is the weighted mean of the σ_{PH}^{-2}'s obtained under Uniform and Natural censoring. Therefore, if $n_{PH}(U)$ and $n_{PH}(N)$ denote the sample sizes obtained under Uniform and Natural censoring, then simple calculations show that

$$n_{PH} = \frac{1}{\frac{\theta}{n_{PH}(U)} + \frac{1-\theta}{n_{PH}(N)}},$$

i.e., n_{PH} actually becomes the weighted harmonic mean of the sample sizes obtained under Uniform and Natural censoring.

If we treat θ to be unknown and include it as a parameter in the model, the likelihood (5.3.1) of the data would be modified as

$$L(\lambda_A, \lambda_B, \theta) \propto \prod_{J=A,B}\prod_{i=1}^{n_J}\{f_J(x_{Ji})^{\Delta_{Ji}} S_J(x_{Ji})^{1-\Delta_{Ji}} f_C(x_{Ji})^{1-\Delta_{Ji}} S_C(x_{Ji})^{\Delta_{Ji}}\},$$

where $f_C(x_{Ji}) = \begin{cases} \frac{\theta}{L} & \text{if } 0 \le x_{Ji} < L, \\ 1-\theta & \text{if } x_{Ji} = L \\ 0 & \text{otherwise} \end{cases}$,

$0 < \theta < 1$ is the probability distribution of the common censoring distribution C, and $S_C(x_{Ji}) = P(C \ge t)$ is its survival function.

Since the survival and the censoring distributions do not share parameters, and the parameter of interest VE_H does not depend on θ, the MLE of λ_A, λ_B, and VE_H will be the same as in Section 5.3.1. Evidently, the subsequent calculations and the asymptotic distribution of γ will also remain unaffected by the change in likelihood. Therefore, the sample size formula n_{exp} will stay the same as derived in the case of known θ. Clearly, the same statement can be made regarding n_{exp} and n_{PH} as well. But all the sample sizes depend on the unknown value of θ along with other parameters, and we have to guess its value to get an insight about the sample size before the beginning of the trial. As usual, the guess may be from the estimates of a previous phase (like Phase II estimates used to derive sample sizes in Phase III) or from some previously conducted trial.

The implementation part will, however, become complicated as θ needs to be estimated

adaptively at every stage along with the other parameters. Obtaining the MLE of θ from the likelihood is difficult and would require numerical solution. The asymptotic properties of RAR designs, SMLE, DBCD, and ERADE will change and need to be studied carefully.

5.8 Concluding Remarks

This article aims to provide an insight about the required sample size in vaccine trials after identifying the time-to-event nature of their datasets. All the sample sizes in this chapter have been computed by considering the asymptotic distribution of the various estimates of γ, which is likely to hold true for the large sample sizes used in trials. The original Binomial model used in most vaccine trials is a misspecified one, and we have shown that it gives misleading sample sizes (see Table 5.3). The survival model, on the other hand, correctly computes the required sample sizes while accounting for factors like loss of information due to censoring, early termination of study, etc. Even the optimum allocations like Neyman and RSIHR worked well and achieved their respective objectives. However, throughout the chapter we assumed that every individual is followed for the same period L unless censored before. This may not hold true in real-life trials, as individuals may get recruited at different time points, resulting in different follow-up periods. If S denotes the length of the entire study, and one assumes that the individuals are accrued uniformly between time 0 and A ($A < S$), then the follow-up period has often been set to $S - A$ (Zhang & Rosenberger, 2007; Zhang & Quan, 2009). Therefore, we can obtain crude estimates of the sample sizes by simply setting $S = L - A$ and proceeding as previously. Nevertheless, the sample size determination problem by incorporating unequal (or random) follow-up times may be explored in the future.

One major issue with the sample size determination problem is that the calculations are done before the beginning of the study, and we have very limited information about the behavior of the upcoming observations from Phase II or previously conducted trials. The problem is more pronounced in the case of survival data, as we need to make assumptions about the distributions of the survival and censoring time with that scarce information. The semi-parametric approach described in Section 5.3.3 partially solves the problem as it does not require any distributional assumptions on the survival times. The parametric approach, the likes of which are more widely available in literature (Zhang & Rosenberger, 2007; Biswas et al., 2016; Das & Bhattacharya, 2022), has been explored as well. The two methods have been compared and shown to give identical results, making the semi-parametric approach more appealing because of its easier calculations and flexibility to any distribution.

Another important aspect of the sample size determination problem — the implementation part — has also been discussed briefly for the parametric approach only. The

implementation in the case of the semi-parametric approach has not been explored, as the problem possesses a lot of difficulties, and the problem has to be treated separately. For example, it would involve estimating the baseline hazard function and checking whether assumption (5.6.1) holds true for the PH model. Nevertheless, our primary objective was fulfilled, which was to give an initial estimate of the sample size before the start of the trial.

Although other forms of the censoring distribution may be taken into consideration, form (5.2.1) is more suitable for realistic scenarios. However, the quantity θ is unknown, and it is somewhat related to the loss-to-follow-up/dropout rate, and it is difficult to get an idea beforehand about the proportion of people that will leave the study prematurely. We have considered the two extreme cases: $\theta = 0$ and $\theta = 1$, and the actual sample sizes are expected to lie in between. In reality, it is expected that θ is small, and the sample sizes are closer to the ones obtained under Natural censoring. A value of θ close to 1 is undesirable, and it is due to this reason that uniform censoring has been seen to overestimate the sample sizes. In contrast, Natural censoring computes the sample sizes under ideal conditions, which provides a preliminary estimate about the minimum number of individuals to recruit. However, it is important to remember that we might require more participants later in the study to offset any early subject withdrawals.

Chapter 6

Future Research Directions

"What we know is a drop, what we don't know is an ocean."

— Commonly attributed to Isaac Newton; origin unverified.

All the chapters in this thesis have certain future aspects that were left unattended. We briefly discuss these future directions one by one in this final concluding chapter.

6.1 Future Aspects of Chapter 2

The ESSs derived in Chapter 2 served our purpose for the time being, but they lacked certain important characteristics. For example, even though we have compared the Prior ESS and the Average Posterior ESS $(-n)$, the current methods do not satisfy the predictive consistency criterion. Development of methods that satisfy the predictive consistency criterion in the case of mixed response models that provide reliable values in case of non-informative priors requires further research. Another interesting future direction is *to incorporate covariates* into our models, as the proposed methods did not account for any heterogeneity that may be present in the form of external covariates.

6.2 Future Aspects of Chapter 3

In Section 3.4.1, we introduced the conditional binomial model and showed that the value of θ under the null hypothesis (3.4.4), $\theta_0 = \frac{1}{1+c}$, becomes a function of the allocation ratio c. Therefore, the hypothesis changes with the value of c. The methods work perfectly when c is completely specified, but the problem changes dratically when c has to be optimally determined. Therefore, we may explore developing Neyman and RSIHR-like allocations in such cases.

6.3 Future Aspects of Chapter 4

We already mentioned that Chapter 4 is a Bayesian alternative to Chapter 3. However, one important aspect of Chapter 3 left unattended in Chapter 4 is the determination of sample size by fixing the power of the relevant test. In fact, many real-life trials, including the recently developed Bayesian COVID-19 trials (Harrell & Lindsell, 2020), often fix the power of the relevant test beforehand instead of the precision of the study. Since the two approaches are asymptotically equivalent in the frequentist case, it is expected that they will behave similarly in the Bayesian case as well, allowing our methods to be applied. Nevertheless, developing Bayesian sample size by fixing power may still serve as an intriguing future problem.

6.4 Future Aspects of Chapter 5

A potential direction from Chapter 5 is a thorough investigation of the implementation part of the sample size determination problem. We have already mentioned in Section 5.8 that the implementation in the case of the semi-parametric approach has a lot of impediments, and we cannot simply extend the RAR techniques, SMLE, DBCD, and ERADE, to achieve the allocation proportions ρ^{PH}_{Neyman} and ρ^{PH}_{RSIHR}. All the underlying presumptions of the RAR techniques need to be satisfied, and the variabilities of the procedures need to be studied theoretically as well as with the help of simulation. Additionally, if we consider the mixture censoring distribution and introduce a new parameter θ in the model, then the algorithm as well as the asymptotic properties of the RAR techniques need to be modified in the light of the new model.

The most important future work in the context of Chapter 5 is perhaps the treatment of the problem when VE_H is time-dependent and the efficacy changes over the course of time. However, when we remove the assumption $\gamma(t) = \gamma$ for all t, the problem becomes more complicated as we need to estimate the function $VE_H(t)$ instead of the single-valued parameter VE_H.

Another important avenue of the sample size determination problem that is yet to be extensively researched is the presence of heterogeneity, which occurs when the vaccine has a different effect across various strata of the vaccinated population. The population may be divided into various strata due to some pre-vaccination host condition(s) or a vaccine-related problem. Pre-vaccination host conditions are simply the presence of host-related factors like age, comorbidities, weight, etc., that change the susceptibility of the vaccine. Vaccine-related factors are heterogeneity among the different vaccine doses that affects the way in which the vaccine works in the individuals. For example, vaccines manufactured from two different batches or administered through two different routes,

like nasal aerospray or intramuscular, or applied in two different populations in different geographical regions may show different efficacies when applied to the host. The vaccine efficacy in the presence of such covariates is calculated as the average relative reduction in susceptibility in all the strata combined (Halloran et al., 1992).

In today's world, when adaptive clinical trials are becoming increasingly popular, an interesting addition to the thesis would be the sample size re-estimation problem. It is a kind of adaptive design that allows for modification of the sample size midway through the trial based on interim results. We only concentrated on providing an initial estimate of the sample size, but the methods can be extended to encompass situations that can re-evaluate the estimates based on accumulating data. The ENSEMBLE Phase 3 trial for Janssen's Ad26.COV2.S COVID-19 vaccine initially targeted around $60,000$ participants, which was later on reduced to around $40,000$. The revision was done through a simple protocol amendment, as the COVID incident rates were seen to be higher than those assumed at the time of protocol planning (see `https://clinicaltrials.gov/study/NCT04505722`). Even though no formal statistical sample size re-estimation methods (Bauer & Köhne, 1994; Proschan & Hunsberger, 1995; Jennison & Turnbull, 2015) were followed, this particular instance shows the potential applicability of sample size adaptations in future trials.

Lights off!!

Appendix A

Appendix for Chapter 2

A.1 Sketch of the Derivation of the Posterior Distribution (2.2.2) Mentioned in Section 2.2

If $p(\boldsymbol{\theta})$ denote the prior density as evidenced by (2.2.1), the posterior distribution of $\boldsymbol{\theta}|(y_i, z_i), i = 1, 2, \ldots, n$ is given by,

$$\begin{aligned} p(\boldsymbol{\theta}|(y_i, z_i), i = 1, 2, \ldots, n) &\propto p(\boldsymbol{\theta}) L(\boldsymbol{\theta}|(y_i, z_i), i = 1, 2, \ldots, n) \\ &\propto p^{a+n_1-1}(1-p)^{b+n_0-1} \exp\Bigg\{ -\frac{1}{2\sigma^2}\sum_{i=1}^{n}(z_i - (1-y_i)\mu_0 - y_i\mu_1)^2 \\ &\qquad -\frac{1}{2(1-\rho^2)}\{(\frac{\mu_0-\gamma_0}{\sigma_0})^2 \\ &\qquad +(\frac{\mu_1-\gamma_1}{\sigma_1})^2 - 2\rho(\frac{\mu_0-\gamma_0}{\sigma_0})(\frac{\mu_1-\gamma_1}{\sigma_1})\Bigg\}. \end{aligned} \tag{A.1.1}$$

It is quite transparent from equation (A.1.1) that $p|\{(y_i, z_i), i = 1, 2, \ldots, n\} \sim Beta(a+n_1, b+n_0)$ and is independent of $\boldsymbol{\mu}|\{(y_i, z_i), i = 1, 2, \ldots, n\}$. The exponent term of (A.1.1) can be simplified (up to some constant) as

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(z_i-(1-y_i)\mu_0-y_i\mu_1)^2+\frac{1}{(1-\rho^2)}\left\{\left(\frac{\mu_0-\gamma_0}{\sigma_0}\right)^2+\left(\frac{\mu_1-\gamma_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{\mu_0-\gamma_0}{\sigma_0}\right)\left(\frac{\mu_1-\gamma_1}{\sigma_1}\right)\right\}$$
$$=\mu_0^2\left(\frac{n_0}{\sigma^2}+\frac{1}{\sigma_0^2(1-\rho^2)}\right)+\mu_1^2\left(\frac{n_1}{\sigma^2}+\frac{1}{\sigma_1^2(1-\rho^2)}\right) - 2\mu_0\left(\frac{\sum_{i=1}^{n}z_i(1-y_i)}{\sigma^2}+\frac{\gamma_0}{\sigma_0^2(1-\rho^2)}-\frac{\rho\gamma_1}{\sigma_0\sigma_1(1-\rho^2)}\right)$$
$$-2\mu_1\left(\frac{\sum_{i=1}^{n}z_iy_i}{\sigma^2}+\frac{\gamma_1}{\sigma_0^2(1-\rho^2)}-\frac{\rho\gamma_0}{\sigma_0\sigma_1(1-\rho^2)}\right)-2\mu_0\mu_1\frac{\rho}{\sigma_0\sigma_1(1-\rho^2)}. \tag{A.1.2}$$

We see that the exponent is a quadratic polynomial in μ_0 and μ_1. Therefore, the posterior distribution of $\boldsymbol{\mu}$ will be bivariate normal, i.e. for some $\boldsymbol{\gamma}^* = \begin{pmatrix}\gamma_0^*\\ \gamma_1^*\end{pmatrix}$ and $\Sigma^* = \begin{bmatrix}\sigma_0^{*^2} & \rho^*\sigma_0^*\sigma_1^*\\ \rho^*\sigma_0^*\sigma_1^* & \sigma_1^{*^2}\end{bmatrix}$,

$$\boldsymbol{\mu}|\{(y_i,z_i), i=1,2,\ldots,n\}\sim N_2(\boldsymbol{\gamma}^*,\Sigma^*).$$

Equating the coefficients of the exponent of $N_2(\boldsymbol{\gamma}^*,\Sigma^*)$ density with the RHS of equation (A.1.2), we get

$$\frac{1}{\sigma_0^{*2}(1-\rho^{*2})}=\frac{n_0}{\sigma^2}+\frac{1}{\sigma_0^2(1-\rho^2)}, \tag{A.1.3}$$
$$\frac{1}{\sigma_1^{*2}(1-\rho^{*2})}=\frac{n_1}{\sigma^2}+\frac{1}{\sigma_1^2(1-\rho^2)}, \tag{A.1.4}$$
$$\frac{\gamma_0^*}{\sigma_0^{*2}(1-\rho^{*2})}-\frac{\rho^*\gamma_1^*}{\sigma_0^*\sigma_1^*(1-\rho^{*2})}=\frac{\sum_{i=1}^{n}z_i(1-y_i)}{\sigma^2}+\frac{\gamma_0}{\sigma_0^2(1-\rho^2)}-\frac{\rho\gamma_1}{\sigma_0\sigma_1(1-\rho^2)}, \tag{A.1.5}$$
$$\frac{\gamma_1^*}{\sigma_1^{*2}(1-\rho^{*2})}-\frac{\rho^*\gamma_0^*}{\sigma_0^*\sigma_1^*(1-\rho^{*2})}=\frac{\sum_{i=1}^{n}z_iy_i}{\sigma^2}+\frac{\gamma_1}{\sigma_0^2(1-\rho^2)}-\frac{\rho\gamma_0}{\sigma_0\sigma_1(1-\rho^2)}, \tag{A.1.6}$$
$$\text{and, } \frac{\rho^*}{\sigma_0^*\sigma_1^*(1-\rho^{*2})}=\frac{\rho}{\sigma_0\sigma_1(1-\rho^2)}. \tag{A.1.7}$$

Solving the set of equations (A.1.3) to (A.1.7) for γ_0^*, γ_1^*, σ_0^{*2}, σ_1^{*2} and ρ^*, we get the posterior parameters as mentioned in (2.2.2). □

A.2 Simulation Protocol

The pseudocode for replicating columns 3 to 6 of Tables 2.6 and 2.8 is given in Algorithms 1 and 2 respectively. The pseudocode for Table 2.9 is almost identical to Algorithm 1, and can be obtained along the same lines. Column 2 of the aforementioned Tables can be obtained using formulas mentioned in Tables 2.1 and 2.4.

Algorithm 1: Finding the Average Posterior ESS for model (1.6.1) and prior (2.2.1)

Input : $0 < p < 1$, $\boldsymbol{\mu} \in \mathbb{R}^2$, $\sigma^2 > 0$, $a > 0$, $b > 0$, $\boldsymbol{\gamma} \in \mathbb{R}^2$, Σ (positive definite matrix), $n \in \mathbb{N}$, $n_{iter} \in \mathbb{N}$, `method`$\in \{VR, PR, MTM, MTM.P, ELIR\}$ (type of ESS)

Output : Average Posterior ES, Standard Error

Define array `Posterior` ESS of length n_{iter};
for $iter = 1, 2, \ldots, n_{iter}$ **do**
  Generate n observations $\{(y_i, z_i), i = 1, 2, \ldots, n\}$ following model (1.6.1);
  Compute the posterior parameters a^*, b^*, γ^*, Σ^* as in equation (2.2.2);
  Find the Posterior ESS (specified `method`) by replacing a, b, γ, Σ in Table 2.1 by a^*, b^*, γ^*, Σ^* ;
  Store the value of Posterior ESS in array `Posterior` ESS;
return Average Posterior ESS = mean(`Posterior` ESS) and Standard Error = sd(`Posterior` ESS);

Algorithm 2: Finding the Average Posterior ESS for model (1.6.1) and prior (2.4.1)

Input : $0 < p < 1$, $\boldsymbol{\mu} \in \mathbb{R}^2$, $\sigma^2 > 0$, $a > 0$, $b > 0$, $d \in \mathbb{N}$, $\boldsymbol{\gamma} \in \mathbb{R}^2$, Σ (positive definite matrix), $n \in \mathbb{N}$, $n_{iter} \in \mathbb{N}$, $N \in \mathbb{N}$ (number of posterior samples to be generated), `method`$\in \{VR, PR, MTM, MTM.P, ELIR\}$ (type of ESS)

Output : Average Posterior ES, Standard Error

Define array `Posterior` ESS of length n_{iter};
for $iter = 1, 2, \ldots, n_{iter}$ **do**
  Generate n observations $\{(y_i, z_i), i = 1, 2, \ldots, n\}$ following model (1.6.1);
  Generate N observations from the posterior distribution (2.4.2) (Use Metropolis Hastings algorithm with proposal density Beta Bivariate normal);
  Find the Posterior ESS (simulated) using the generated samples;
  Store the value of Posterior ESS in array `Posterior` ESS;
return Average Posterior ESS = mean(`Posterior` ESS) and Standard Error = sd(`Posterior` ESS);

Chosen value of the inputs:

$p = 0.6$, $\mu = \begin{pmatrix} 0 \\ 1.5 \end{pmatrix}$, $\sigma^2 = 64$, $a = 100$, $b = 75$, $\boldsymbol{\gamma} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$, $\Sigma = \begin{bmatrix} 4 & 5.4 \\ 5.4 & 9 \end{bmatrix}$, $d = 5$, $n = 10/100/1000/10,000$ (Output generated for each n), $n_{iter} = 10000$, $N = 1000$ and `method`=VR/PR/MTM/MTM.P/ELIR.

Appendix B

Appendix for Chapter 3

B.1 Simulation protocol followed for generating Table 3.10

In this section, we provide certain algorithms so that the results obtained in columns 4 to 9 of Table 3.10 can be reproduced. Columns 1 and 2 can be obtained by plugging in the appropriate input values in formulas (3.4.3) and (3.4.2). As the simulations have been done both in the presence and absence of drift, we have a pseudocode for each. In the beginning of the algorithms, all the necessary input variables have been mentioned along with their possible values. The particular input values chosen by us have also been reported.

Algorithm 3: Pseudocode for without drift simulation

Input : $VE \in (0,1)$, $ARU/p_A \in (0,1)$, $p_B = (1-VE)p_A$, $\alpha \in (0,1)$, `desired width`> 0, $n_{iter} \in \mathbb{N}$ (number of iterations), $n_{seed} \in \mathbb{N}$, `allocation type` $\in$ {Equal, Double, Neyman, RSIHR} (specifies the targeted allocation), `RAR Technique` $\in$ {1:1, 1:2, SMLE, DBCD, ERADE} (defaults to 1:1 when `allocation type` = Equal, and 1:2 when `allocation type` = Double), $\alpha_1 > 0$ (only needed when `RAR Technique`= SMLE or DBCD; defaults to 0 in SMLE), $\alpha_2 > 0$ (only needed if `RAR technique`=ERADE).

Output : Expected Sample Size (Total), Expected sample size (Vaccine) and Simulated power

for *iter = 1,2,...,n_{iter}* **do**
- Define arrays n^{array}, n_B^{array} and `rejection indicator`, each of length n_{iter} ;
- Allocate first $2n_{seed}$ patients to vaccine or placebo using (1:1) randomization;
- Simulate responses for first $2n_{seed}$ patients – Generate observation from $Bin(1, p_A)$ if allocated to placebo and from $Bin(1, p_B)$ if allocated to vaccine (Store the allocation and response histories at all steps);
- Calculate $\widehat{p}_A$, $\widehat{p}_B$, width of the $100(1-\alpha)\%$ CI for VE (W) and proportion allocated to placebo till now ($\widehat{\rho}$);
- Set $i = 2n_{seed}$;
- **while** $W >$ *`desired width`* **do**
  - i=i+1;
  - Allocate patient i to placebo with probability $g(\widehat{\rho}, \rho(\widehat{p}_A, \widehat{p}_B))$ [The function g depends on input `RAR technique`. $g(x,y) = \frac{1}{2}$ if `RAR Technique` = 1:1; $g(x,y) = \frac{1}{3}$ if `RAR Technique` = 1:2; $g(x,y) = \begin{cases} \alpha_2 y & \text{if } x > y \\ y & \text{if } x = y \\ 1-\alpha_2(1-y) & \text{if } x < y \end{cases}$ if `RAR Technique` = ERADE; and $g(x,y)$ as in equation (3.7.1) if `RAR Technique` = SMLE or DBCD. Note that, $\alpha_1 = 0$ for SMLE. $\rho(p_A, p_B)$ depend on `allocation type` (see equations (3.3.2) and (3.3.4)).];
  - Simulate response for patient i – Generate observation from $Bin(1, p_A)$ if allocated to placebo and from $Bin(1, p_B)$ if allocated to vaccine;
  - Update $\widehat{p}_A$, $\widehat{p}_B$, W and $\widehat{\rho}$;
- Perform the level α test for $H_0 : VE \leq 0$ vs. $H_1 : VE > 0$;
- Store 1 inside `rejection indicator` if H_0 is rejected, and 0 otherwise;
- Store the value of i inside n_{array} and the number of allocations to vaccine inside n_B^{array};

return Expected Sample Size (Total) = mean(n_{array}), Expected sample size (Vaccine) = mean(n_B^{array}) and Simulated power = mean(`rejection indicator`)

Algorithm 4: Pseudocode for with drift simulation

Input : $VE \in (0,1)$, $ARU/p_A \in (0,1)$, $p_B = (1-VE)p_A$, $\alpha \in (0,1)$, `desired width`> 0, $n_{iter} \in \mathbb{N}$ (number of iterations), $n_{seed} \in \mathbb{N}$, `drift` > 0, `allocation type` $\in$ {Equal, Double, Neyman, RSIHR} (specifies the targeted allocation), `RAR Technique` $\in$ {1:1, 1:2, SMLE, DBCD, ERADE} (defaults to 1:1 when `allocation type` = Equal, and 1:2 when `allocation type` = Double), $\alpha_1 > 0$ (only needed when `RAR Technique`= SMLE or DBCD; defaults to 0 in SMLE), $\alpha_2 > 0$ (only needed if `RAR technique`=ERADE).

Output : Expected Sample Size (Total), Expected sample size (Vaccine) and Simulated power

for *iter = 1,2,...,n_{iter}* **do**
- Define arrays n^{array}, n_B^{array} and `rejection indicator`, each of length n_{iter} ;
- Allocate first $2n_{seed}$ patients to vaccine or placebo using (1:1) randomization;
- Simulate responses for first $2n_{seed}$ patients – For patient i, generate observation from $Bin(1, p_A\texttt{drift}^i)$ if allocated to placebo and from $Bin(1, p_B\texttt{drift}^i)$ if allocated to vaccine, $i = 1, 2, \ldots, 2n_{seed}$ (Store the allocation and response histories at all steps);
- Calculate $\widehat{p}_A$, $\widehat{p}_B$, width of the $100(1-\alpha)\%$ CI for VE (W) and proportion allocated to placebo till now $(\widehat{\rho})$;
- Set $i = 2n_{seed}$;
- **while** $W >$ *`desired width`* **do**
  - i=i+1;
  - Allocate patient i to placebo with probability $g(\widehat{\rho}, \rho(\widehat{p}_A, \widehat{p}_B))$ [The functions g and ρ are same as in with drift simulation];
  - Simulate response for patient i – Generate observation from $Bin(1, p_A\texttt{drift}^i)$ if allocated to placebo and from $Bin(1, p_B\texttt{drift}^i)$ if allocated to vaccine;
  - Update $\widehat{p}_A$, $\widehat{p}_B$, W and $\widehat{\rho}$;
- Perform the level α test for $H_0 : VE \leq 0$ vs. $H_1 : VE > 0$;
- Store 1 inside `rejection indicator` if H_0 is rejected, and 0 otherwise;
- Store the value of i inside n_{array} and the number of allocations to vaccine inside n_B^{array};

return Expected Sample Size (Total) = mean(n_{array}), Expected sample size (Vaccine) = mean(n_B^{array}) and Simulated power = mean(`rejection indicator`)

Chosen value of the inputs:

$VE = 0.4$, $ARU/p_A = 0.1$, $p_B = 0.06$, $\alpha = 0.05$, `desired width` $= 0.5$, $n_{iter} = 1000$, $n_{seed} = 5$, `drift` $= 1.001$, `allocation type` = Equal/Double/Neyman/RSIHR, `RAR Technique` = 1:1/1:2/SMLE/DBCD/ERADE, $\alpha_1 = 2$ (when `RAR Technique` = DBCD), $\alpha_2 = 0.5$.

Appendix C

Appendix for Chapter 4

C.1 Derivation of the Asymptotic Posterior Distribution of ψ

The asymptotic distribution will be derived using the Berstien Von-Mises theorem stated below.

Bernstein Von-Mises Theorem Suppose $X_1, X_2, \ldots, X_n$ are iid observations, abbreviated as $\boldsymbol{X_n}$, having a density $f(x_n \mid \theta), \theta \in \Theta \subset \mathbb{R}^p$. Let $\pi(\theta)$ be a prior density and $\pi(\theta \mid \boldsymbol{X_n})$ the posterior density. Let $\tilde{\theta}_n$ be the posterior mode, and $\tilde{I}_n$ be the Fisher information matrix (evaluated at $\tilde{\theta}_n$) defined below. Then, under suitable regularity conditions, for large n, the posterior distribution of θ,

$$\pi(\theta \mid \boldsymbol{X_n}) \overset{a}{\sim} N_p\left(\tilde{\theta}_n, \tilde{I}_n^{-1}\right),$$

where $\tilde{I}_n = \left(-\frac{\partial^2}{\partial\theta_i\theta_j}\log\pi(\theta \mid \boldsymbol{X_n})\right)\Big|_{\theta=\tilde{\theta}_n}$ (Ghosh et al., 2006; Van der Vaart, 1998).

Considering equations (1.3.1) and (4.3.3), we have joint posterior density of p_A and p_B as

$$\pi(p_A, p_B \mid x_{An_A}, x_{Bn_B}) = \frac{1}{B(a_A^*, b_A^*)B(a_B^*, b_B^*)} p_A^{a_A^*-1}(1-p_A)^{b_A^*-1} p_B^{a_B^*-1}(1-p_B)^{b_B^*-1}.$$

Therefore, the joint posterior density of p_A and ψ can be obtained as,

$$\pi(p_A, \psi \mid x_{An_A}, x_{Bn_B}) = \frac{1}{B(a_A^*, b_A^*)B(a_B^*, b_B^*)} p_A^{a_A^*+a_B^*-1}(1-p_A)^{b_A^*-1}\psi^{a_B^*-1}(1-\psi p_A)^{b_B^*-1}.$$

The posterior mode of the above density is given by $(\widetilde{p}_A, \widetilde{\psi})$ where $\widetilde{p}_A = \frac{a_A^*}{a_A^*+b_A^*-1}$, $\widetilde{\psi} = \frac{\widetilde{p}_B}{\widetilde{p}_A}$ and $\widetilde{p}_B = \frac{a_B^*-1}{a_B^*+b_B^*-2}$.

Hence, the Fisher information matrix $\widetilde{I}_n$ is

$$\begin{aligned}\widetilde{I}_n &= \begin{bmatrix} -\frac{\partial^2}{\partial p_A^2}\log\pi(p_A,\psi \mid x_{An_A}, x_{Bn_B}) & -\frac{\partial^2}{\partial p_A \partial \psi}\log\pi(p_A,\psi \mid x_{An_A}, x_{Bn_B}) \\ -\frac{\partial^2}{\partial \psi \partial p_A}\log\pi(p_A,\psi \mid x_{An_A}, x_{Bn_B}) & -\frac{\partial^2}{\partial \psi^2}\log\pi(p_A,\psi \mid x_{An_A}, x_{Bn_B}) \end{bmatrix} \\ &= \begin{bmatrix} \frac{(a_A^*+b_A^*-1)}{\widetilde{p}_A(1-\widetilde{p}_A)} + \frac{(a_B^*+b_B^*-2)\widetilde{p}_B}{\widetilde{p}_A^2(1-\widetilde{p}_B)} & \frac{(a_B^*+b_B^*-2)}{(1-\widetilde{p}_B)} \\ \frac{(a_B^*+b_B^*-2)}{(1-\widetilde{p}_B)} & \frac{(a_B^*+b_B^*-2)\widetilde{p}_A^2}{\widetilde{p}_B(1-\widetilde{p}_B)} \end{bmatrix}.\end{aligned}$$

It is clear that the asymptotic posterior mean of ψ is $\widetilde{\psi}$ and its variance is the $(4,4)^{th}$ element of $\widetilde{I}_n^{-1}$ given by

$$v(\tilde{p_A}, \tilde{p_B}) = \frac{\tilde{p_B}^2(1-\tilde{p_A})}{\tilde{p_A}^3(a_A+b_A+n_A-1)} + \frac{\tilde{p_B}(1-\tilde{p_B})}{\tilde{p_A}^2(a_B+b_B+n_B-2)}.$$

□

C.2 Proof of Lemma 4.5.1

Let us denote the first and second term on the RHS of equation (4.5.1) by $v_1(\tilde{p_A}, \tilde{p_B})$ and $v_2(\tilde{p_A}, \tilde{p_B})$, respectively. Now,

$$\begin{aligned}E_{\boldsymbol{x}}\{v_1(\tilde{p_A}, \tilde{p_B})\} &= \sum_{x_{An_A}=0}^{n_A}\sum_{x_{Bn_B}=0}^{n_B} \frac{\tilde{p_B}^2(1-\tilde{p_A})}{\tilde{p_A}^3(a_A+b_A+n_A-1)} f(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta}) \\ &= (a_A+b_A+n_A-1) \\ &\quad \times \sum_{x_{An_A}=0}^{n_A} \frac{(b_A-1+n_A-x_A)}{(a_A+x_A)^3}\binom{n_A}{x_{An_A}}\frac{Beta(a_A+x_A, n_A+b_A-x_{An_A})}{Beta(a_A,b_A)} \\ &\quad \times \sum_{x_{Bn_B}=0}^{n_B} \frac{(a_B+x_B-1)^2}{(a_B+b_B+n_B-2)^2}\binom{n_B}{x_{Bn_B}}\frac{Beta(a_B+x_B, n_B+b_B-x_{Bn_B})}{Beta(a_B,b_B)} \\ &\leq (a_A+b_A+n_A-1)\sum_{x_{An_A}=0}^{n_A} \frac{(b_A+n_A-x_A)}{(a_A+x_A-1)(a_A+x_A-2)(a_A+x_A-3)} \\ &\quad \times \binom{n_A}{x_{An_A}}\frac{Beta(a_A+x_A, n_A+b_A-x_{An_A})}{Beta(a_A,b_A)} \\ &\quad \times \sum_{x_{Bn_B}=0}^{n_B} \frac{(a_B+x_B)(a_B+x_B+1)}{(a_B+b_B+n_B-2)^2}\binom{n_B}{x_{Bn_B}}\frac{Beta(a_B+x_B, n_B+b_B-x_{Bn_B})}{Beta(a_B,b_B)} \\ &= \frac{1}{(a_A+b_A+n_A-2)}\cdot\frac{(a_A+b_A-1)(a_A+b_A-2)b_A}{(a_A-1)(a_A-2)(a_A-3)}\end{aligned}$$

$$\times \frac{(a_B+b_B+n_B)(a_B+b_B+n_B+1)}{(a_B+b_B+n_B-2)^2}.\frac{a_B(a_B+1)}{(a_B+b_B)(a_B+b_B+1)}$$
$$= \frac{(a_B+b_B+n_B)(a_B+b_B+n_B+1)}{(a_A+b_A+n_A-2)(a_B+b_B+n_B-2)^2}.v_1(\boldsymbol{\eta}). \quad \text{(C.2.1)}$$

Similarly, it can be shown that

$$E_{\boldsymbol{x}}\{v_2(\widetilde{p_A},\widetilde{p_B})\} = \sum_{x_{An_A}=0}^{n_A}\sum_{x_{Bn_B}=0}^{n_B}\frac{\widetilde{p_B}(1-\widetilde{p_B})}{\widetilde{p_A}^2(a_B+b_B+n_B-2)}f(x_{An_A},x_{Bn_B}\mid n_A,n_B,\boldsymbol{\eta})$$
$$\leq \frac{(a_A+b_A+n_A-1)(a_B+b_B+n_B)(a_B+b_B+n_B+1)}{(a_A+b_A+n_A-2)(a_B+b_B+n_B-2)^3}.v_2(\boldsymbol{\eta}). \quad \text{(C.2.2)}$$

Combining (C.2.1) and (C.2.2), we get the first part of the proof. Also, if n_A and n_B are large enough with $\frac{n_A}{n_A+n_B} \longrightarrow \rho$, then

$$\frac{(a_B+b_B+n_B)(a_B+b_B+n_B+1)}{(a_A+b_A+n_A-2)(a_B+b_B+n_B-2)^2}.v_1(\boldsymbol{\eta})$$
$$+\frac{(a_A+b_A+n_A-1)(a_B+b_B+n_B)(a_B+b_B+n_B+1)}{(a_A+b_A+n_A-2)(a_B+b_B+n_B-2)^3}.v_2(\boldsymbol{\eta})$$
$$\approx \frac{v_1(\boldsymbol{\eta})}{(a_A+b_A+n_A-2)}+\frac{v_2(\boldsymbol{\eta})}{(a_B+b_B+n_B-2)},$$

which completes the proof. □

C.3 Simulation protocol

This section outlines the simulation protocol followed in Chapter 4. We provide three algorithms that can be used to reproduce columns 5 and 8 of Tables 4.1 to 4.3 in the chapter. Algorithms 6 and 7 respectively describe finding the ACC and ALC sample sizes under Fixed and Neyman allocation, respectively. The algorithms for ACC'' sample sizes (column 6), RSIHR allocation, and case-control studies are similar and have been omitted for the sake of brevity. The methods for finding the approximate bounds (last column of the tables) involve simple formulas mentioned in Results 4.5.1 and 4.5.2. We mentioned that ACC and ALC sample size computations can be made easier with the help of these bounds. The algorithms consider these bounds as inputs (n_A^{seed}, n_B^{seed}) and search around them to find the actual ACC and ALC sample sizes. However, both algorithms involve finding $acc(n_A, n_B, \boldsymbol{\eta})$ and $alc(n_A, n_B, \boldsymbol{\eta})$, and the pseudocode for finding them using the Monte Carlo technique has been outlined in Algorithm 5. The pseudocode for the asymptotic method has been omitted, as it is simpler and can be obtained from

equations (4.4.8) and (4.4.9).

Algorithm 5: Finding $acc(n_A, n_B, \boldsymbol{\eta})$ (or $alc(n_A, n_B, \boldsymbol{\eta})$ using the Monte Carlo method for HPD interval computation

Input : $n_A \in \mathbb{N}$, $n_B \in \mathbb{N}$, Hyper-parameters $\boldsymbol{\eta} = (a_A, b_A, a_B, b_B)$, $l > 0$, $0 < \alpha < 1$, $N \in \mathbb{N}$ (number of posterior samples used)

Output : $acc(n_A, n_B, \boldsymbol{\eta})$ (or $acc(n_A, n_B, \boldsymbol{\eta})$)

Initialization: $acc(n_A, n_B, \boldsymbol{\eta})$(or $alc(n_A, n_B, \boldsymbol{\eta})$)=0

for $x_{An_A} = 0, 1, 2, \ldots, n_A$ **do**
  for $x_{Bn_B} = 0, 1, \ldots, n_B$ **do**
    Generate N observations from $\psi \mid x_{An_A}, x_{Bn_B}$ (exact distribution);
    Find $\widehat{\alpha}_l^*(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta})$ (or $\widehat{l}_{1-\alpha}^*(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta})$) [see equation (4.4.5)(or equation (4.4.6))];
    Find Beta-Binomial density $f(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta})$ [see equation (4.3.4)];
    $acc(n_A, n_B, \boldsymbol{\eta}) = acc(n_A, n_B, \boldsymbol{\eta}) + \widehat{\alpha}_l^*(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta}) f(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta})$ (or $alc(n_A, n_B, \boldsymbol{\eta}) = alc(n_A, n_B, \boldsymbol{\eta}) + \widehat{l}_{1-\alpha}^*(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta}) f(x_{An_A}, x_{Bn_B} \mid n_A, n_B, \boldsymbol{\eta})$) ;

return $acc(n_A, n_B, \boldsymbol{\eta})$ (or $alc(n_A, n_B, \boldsymbol{\eta})$);

Algorithm 6: Finding ACC(or ALC) sample sizes under Fixed allocation

Input : $l > 0$, $0 < \alpha < 1$, $\boldsymbol{\eta}$, $n_A^{seed} \in \mathbb{N}$, `allocation type`$\in$ {Equal, Double}, `HPD computation type` $\in$ {Asymptotic, Monte Carlo}, $N \in \mathbb{N}$ (needed only when `HPD computation type` = Monte Carlo

Output : ACC (or ALC) sample size (total and vaccine)

if *allocation type* = *Equal* **then**
  c=1;
  [Specifying the fixed allocation ratio]
else
  c=0.5;

Initialization: $n_A = n_A^{seed}$, $n_B = n_A^{seed}/c$

Find $acc(n_A, n_B, \boldsymbol{\eta})$ (or $alc(n_A, n_B, \boldsymbol{\eta})$) [Use Monte Carlo or asymptotic method as specified by argument `HPD Technique`];

if $acc(n_A, n_B, \boldsymbol{\eta}) \geq 1 - \alpha$ *(or* $alc(n_A, n_B, \boldsymbol{\eta}) \leq l$*)* **then**
  Set $n_A = n_A - 10$ and $n_B = n_A/c$;
else
  continue;

while $acc(n_A, n_B, \boldsymbol{\eta}) < 1 - \alpha$ *(or* $alc(n_A, n_B, \boldsymbol{\eta}) > l$*)* **do**
  $n_A = n_A + 1$ and $n_B = n_A/c$;

return $n_A + n_B$, n_B;

Algorithm 7: Finding ACC(or ALC) sample sizes under Neyman allocation

Input : $l > 0$, $0 < \alpha < 1$, $\boldsymbol{\eta}$, $n_A^{seed} \in \mathbb{N}$, $n_B^{seed} \in \mathbb{N}$, `HPD computation type` $\in$ {Asymptotic, Monte Carlo}, $N \in \mathbb{N}$ (needed only when `HPD computation type` = Monte Carlo

Output : ACC (or ALC) sample size (total and vaccine)

Initialization: $n_A = n_A^{seed}$, $n_B = n_B^{seed}$

Find $acc(n_A, n_B, \boldsymbol{\eta})$ (or $alc(n_A, n_B, \boldsymbol{\eta})$);

if $acc(n_A, n_B, \boldsymbol{\eta}) \geq 1 - \alpha$ *(or* $alc(n_A, n_B, \boldsymbol{\eta}) \leq l$*)* **then**

  $n_A = n_A - 1$ and $n_B = n_B - 1$;

else

  continue;

Create a grid of values $\mathbb{G}$ around (n_A, n_B) ;

while $acc(x, y, \boldsymbol{\eta}) < 1 - \alpha$ *(or* $alc(x, y, \boldsymbol{\eta}) > l$*) for all* $(x, y) \in \mathbb{G}$ **do**

  Find the point in $(x', y') \in \mathbb{G}$ such that $(x', y') = \underset{(x,y)\in\mathbb{G}}{\arg\max}\, acc(x, y, \boldsymbol{\eta})$ (or $(x', y') = \underset{(x,y)\in\mathbb{G}}{\arg\min}\, alc(x, y, \boldsymbol{\eta})$) ;

  Set $n_A = x'$ and $n_B = y'$;

  Update set $\mathbb{G}$ by creating grid around (n_A, n_B);

Add more points to $\mathbb{G}$ if needed and identify the set of points $\mathbb{H}(\mathbb{H} \subseteq \mathbb{G})$ where $acc(n_A, n_B, \boldsymbol{\eta})$ just crosses $1 - \alpha$ (or $alc(n_A, n_B, \boldsymbol{\eta})$ just falls below l);

Find the point in $(x', y') \in \mathbb{H}$ such that $(x', y') = \underset{(x,y)\in\mathbb{H}}{\arg\max}\, acc(x, y, \boldsymbol{\eta})$ (or $(x', y') = \underset{(x,y)\in\mathbb{H}}{\arg\min}\, alc(x, y, \boldsymbol{\eta})$);

Set $n_A = x'$ and $n_B = y'$;

return $n_A + n_B$, n_B;

Chosen value of the inputs:

$l = 0.4$, $\alpha = 0.05$, $N = 1000$ (kept same throughout), , `allocation type` = Equal/-Double, `HPD computation type` = Asymptotic/Monte Carlo.

The arguments $\boldsymbol{\eta} = (a_A, b_A, a_B, b_B)$ are as provided in columns 1 to 4 and (n_A^{seed}, n_B^{seed}) as in the last column (Approximate Bound) of Tables 4.1 to 4.3.

Appendix D

Appendix for Chapter 5

D.1 Derivation of e_J, e_{1J}, e_{2J}, $J = A, B$ Mentioned in Section 5.3.2

It has already been mentioned that

$$e_J = Pr(\Delta_J = 1) = Pr(T_J < C),$$

where $T_J \sim Weibull(\lambda_J, p)$ and $C \sim \delta(L)$ or $C \sim U[0, L]$ according as we assume Natural or Uniform censoring. Hence

$$e_J = \begin{cases} 1 - e^{-\lambda_J L^p} & \text{under Natural censoring} \\ 1 - \frac{1}{pL\lambda^{\frac{1}{p}}}\gamma(\frac{1}{p}, L^P) & \text{under Uniform censoring,} \end{cases}$$

where $\gamma(s, x) = \int_0^x y^{s-1}e^{-y}dy$ is the lower incomplete gamma function. The values $\gamma(1/p, L^p)$ for various p and L have been obtained numerically using R.

Similarly, $e_{1J} = E(X_J^p \log X_J)$ and $e_{2J} = E(X_J^p(\log X_J)^2)$ where $X_J = min(X_J, C), J = A, B$. If g is a function such that $E(|g(T_J)| < \infty)$, then $E(g(X_J))$ can be expressed as

$$\begin{aligned} E(g(X_J)) &= E(g(T_J)|T_J \le C)P(T_J \le C) + E(g(C)|T_J > C)P(T > C) \\ &= \int_0^L g(t)S_C(t)dF_J(t) + \int_0^L g(c)S_J(c)dF_C(c) \\ &= \begin{cases} \int_0^L g(t)p\lambda_J t^{p-1}e^{-\lambda_J t^p}dt + g(L)e^{-\lambda_J L^p} & \text{if } C \sim \delta(L) \\ \int_0^L g(t)p\lambda_J t^{p-1}e^{-\lambda_J t^p}(1 - \frac{t}{L})dt + \int_0^L g(c)e^{-\lambda_J c^p}dc & \text{if } C \sim U[0, L]. \end{cases} \end{aligned}$$

Putting $g(x) = x^p \log x$ and $g(x) = x^p (\log x)^2$ we get,

$$e_{1J} = \begin{cases} \frac{\lambda}{p} \int_0^{L^p} z \log z e^{-\lambda z} dz + L^p \log L e^{\lambda L^p} & \text{if } C \sim \delta(L) \\ \frac{\lambda}{p} \int_0^{L^p} z \log z e^{-\lambda z} dz - \frac{\lambda}{pL} \int_0^{L^p} z^{1+\frac{1}{p}} \log z e^{-\lambda z} dz & \\ \qquad + \frac{1}{p^2 L} \int_0^{L^p} z^{\frac{1}{p}} \log z e^{-\lambda z} dz & \text{if } C \sim U[0, L]. \end{cases}$$

and,

$$e_{2J} = \begin{cases} \frac{\lambda}{p^2} \int_0^{L^p} z (\log z)^2 e^{-\lambda z} dz + L^p (\log L)^2 e^{\lambda L^p} & \text{if } C \sim \delta(L) \\ \frac{\lambda}{p^2} \int_0^{L^p} z (\log z)^2 e^{-\lambda z} dz - \frac{\lambda}{p^2 L} \int_0^{L^p} z^{1+\frac{1}{p}} (\log z)^2 e^{-\lambda z} dz & \\ \qquad + \frac{1}{p^3 L} \int_0^{L^p} z^{\frac{1}{p}} (\log z)^2 e^{-\lambda z} dz & \text{if } C \sim U[0, L]. \end{cases}$$

The integrals of the type $\int_0^{L^p} z^a (\log z)^b e^{-\lambda z} dz$ for different a and b have been obtained numerically using the `intergrate` function in R. □

D.2 Derivation of the Asymptotic Distribution of $\widehat{\gamma}_{PH}$ Mentioned in Section 5.3.3

We revisit some results and notations from Tsiatis (1981) in order to derive the asymptotic distribution.

Suppose T and C respectively denote the true survival time and time to censoring of an individual. T and C are independent given the covariate Z, which is a single covariate with pdf $f(z)$. For a pre-specified time T_0, it is assumed that $C \leq T_0 < \infty$. The observable time until death or censoring and the censoring indicator are respectively denoted by X and Δ respectively. That is,

$$X = min(T, C) \text{ and } \Delta = \begin{cases} 1 & \text{if } T \leq C \\ 0 & \text{if } T > C. \end{cases}$$

If $g(z)$ is a continuous function of z and I_A denote the indicator function of the event A then,

$$E(g(z), t) = E\{g(Z) I_{\{X \geq t\}}\} = E\{g(Z) P(X \geq t | Z)\}.$$

The Cox PH model assumes that the hazard function for an individual with covariate z is

$$\lambda(t|z) = \lambda_0(t) e^{\gamma z},$$

where γ is the regression co-efficient and $\lambda_0(t)$ is the baseline hazard. If $\widehat{\gamma}_{PH}$ denote the Cox's estimate for γ, obtained by maximizing the corresponding partial likelihood, then Theorem 3.2 of Tsiatis (1981) states that

$$\sqrt{n}(\widehat{\gamma}_{PH} - \gamma) \xrightarrow{d} N(0, \sigma^2_{PH}),$$

where $\sigma^2_{PH} = [\int_0^{T_0} -dQVar(z|R(t))]^{-1}$. Here $Q(t) = P(X \geq t, \Delta = 1)$ is the probability of surviving until time t without being censored and eventually dying before being censored and

$$Var(z|R(t)) = \frac{E(z^2 e^{\gamma z}, t)}{E(e^{\gamma z}, t)} - \left(\frac{E(z e^{\gamma z}, t)}{E(e^{\gamma z}, t)}\right)^2.$$

In our case, we have $T_0 = L$ and the covariate Z is a binary random variable with $P(Z = 1) = 1 - \rho$, and

$$E(g(z), t) = \rho g(0) H_A(t) + (1 - \rho) g(1) h_B(t),$$

where $H_A(t) = P(X \geq t|Z = 0)$ and $H_B(t) = P(X \geq t|Z = 1)$.

If $S_C(t) = P(C \geq t)$ denote the survival function of the censoring distribution C, then

$$H_X(t) = S_X(t) S_C(t), X = A, B. \tag{D.2.1}$$

A routine simplification shows that

$$Var(z|R(t)) = \frac{\rho(1 - \rho) H_A(t) e^{\gamma} H_B(t)}{(\rho H_A(t) + (1 - \rho) e^{\gamma} H_B(t))^2}.$$

Again,

$$\begin{aligned} Q(t) &= P(X \geq t, \Delta = 1) \\ &= \int_t^L \{\rho \lambda_A(x) H_A(x) + (1 - \rho) \lambda_B(x) H_B(x)\} dx \\ \implies \frac{d}{dt} Q(t) &= -\rho \lambda_A(t) H_A(t) - (1 - \rho) \lambda_B(t) H_B(t). \end{aligned}$$

Therefore, the asymptotic variance σ^2_{PH} satisfies

$$\begin{aligned} \sigma^{-2}_{PH} &= \int_0^L -dQVar(z|R(t)) \\ &= \int_0^L \frac{\rho(1 - \rho) H_A(t) e^{\gamma} H_B(t) \lambda_A(t)}{(\rho H_A(t) + (1 - \rho) e^{\gamma} H_B(t))} dt \\ &= \int_0^L \frac{\rho(1 - \rho) e^{\gamma} S_A(t) S_C(t) \lambda_A(t)}{(\rho S_A(t)^{1 - e^{\gamma}} + (1 - \rho) e^{\gamma}} dt \ [\because \text{equation (D.2.1) holds and } S_B(t) = (S_A(t))^{e^{\gamma}}] \end{aligned}$$

$$= \int_0^L \frac{\rho(1-\rho)(1-VE)f_A(t)S_C(t)}{(1-\rho)(1-VE)+\rho S_A(t)^{VE}}dt \text{ [where } f_A(t) \text{ is the pdf of } T_A\text{]}$$

$$= \int_{S_A(L)}^1 \frac{\rho(1-\rho)(1-VE)S_C(S_A^{-1}(y))}{(1-\rho)(1-VE)+\rho y^{VE}}dy \text{ [Putting } y = S_A(t)\text{]}$$

$$= \begin{cases} \int_{S_A(L)}^1 \frac{\rho(1-\rho)(1-VE)}{(1-\rho)(1-VE)+\rho y^{VE}}dy & \text{under Natural censoring} \\ \int_{S_A(L)}^1 \frac{\rho(1-\rho)(1-VE)(1-\frac{S_A^{-1}(y)}{L})}{(1-\rho)(1-VE)+\rho y^{VE}}dy & \text{under Uniform censoring.} \end{cases}$$

□

D.3 Simulation protocol followed in Tables 5.9 and 5.10

The algorithm for generating columns 6 to 8 of Tables 5.9 and 5.10. Columns 2 to 5 can be reproduced using by plugging in the values of λ_A, λ_B and L in equations (5.3.5) and (5.6.2).

Algorithm 8: Finding the simulated values of $E\left(\frac{N_A(n)}{n}\right)$ and $nVar\left(\frac{N_A(n)}{n}\right)$

Input : $\lambda_A \in (0,\infty)$, $VE \in (0,1)$, $\lambda_B = (1-VE) * \lambda_A$, $L \in (0,\infty)$, $\alpha \in (0,1)$, $n \in \mathbb{N}$, $n_{iter} \in \mathbb{N}$, $n_{seed} \in \mathbb{N}$, `censoring` $\in$ {Natural, Uniform} (the censoring mechanism used), `RAR Technique` $\in$ {SMLE, DBCD, ERADE}, `allocation type` $\in$ {Neyman, RSIHR} (allocation targeted through RAR)

Output : $E\left(\frac{N_A(n)}{n}\right)$, $nVar\left(\frac{N_A(n)}{n}\right)$

1 Define array `allocation proportion` of length n_{iter};
2 **for** $iter = 1, 2, \ldots, n_{iter}$ **do**
3 Allocate first $2n_{seed}$ patients to vaccine or placebo using 1:1 randomization ;
4 Generate responses for first $2n_{seed}$ patients based on whether they are allocated to placebo or vaccine. The responses are generated from exponential distribution and censored according to the censoring mechanism considered;
5 Calculate $\widehat{\lambda}_A$, $\widehat{\lambda}_B$ and proportion allocated to placebo till now ($\widehat{\rho}$);
6 **for** $i = 2n_{seed}+1, \ldots, n$ **do**
7 Allocate patient i to placebo with probability $g(\widehat{\rho}, \rho(\widehat{\lambda}_A, \widehat{\lambda}_B))$ [The function g depends on input `RAR technique` and is same as used in Algorithm 3. ρ depends on `allocation type` (see equations (5.3.7) and (5.3.9))];
8 Simulate response for patient i;
9 Update $\widehat{\lambda}_A$, $\widehat{\lambda}_B$ and $\widehat{\rho}$;
10 Store $\widehat{\rho}$ inside `allocation proportion` array ;
11 return mean(`allocation proportion`) and n*variance(`allocation proportion`) ;

Chosen value of the inputs: $\lambda_A = 0.4$, $VE_H = 0.7$, $\lambda_B = 0.12$, $L = 10$, $\alpha = 0.05$, $n = 150/200/300/500$, $n_{iter} = 10,000$, $n_{seed} = 5$, `censoring` = Natural/Uniform, `RAR Technique` = SMLE/DBCD/ERADE, `allocation type` = Neyman/RSIHR.

Bibliography

Adcock, C. J. (1987). A Bayesian approach to calculating sample sizes for multinomial sampling. *J R Stat Soc Ser D Stat*, 36(2-3), 155–159.

Adcock, C. J. (1988). A bayesian approach to calculating sample sizes. *J R Stat Soc Ser D Stat*, 37.

Adcock, C. J. (1997). Sample size determination: a review. *J R Stat Soc Ser D Stat*, 46(2), 261–283.

Agnandji, S. T., Huttner, A., Zinser, M. E., et al. (2016). Phase 1 trials of rVSV Ebola vaccine in Africa and Europe. *N Engl J Med*, 374(17), 1647–1660.

Ahnn, S. & Anderson, S. J. (1995). Sample size determination for comparing more than two survival distributions. *Stat Med*, 14(20), 2273–2282.

Andersen, P. K. & Gill, R. D. (1982). Cox's regression model for counting processes: a large sample study. *Ann Stat*, 10(4), 1100–1120.

Bai, Z. D., Hu, F., & Rosenberger, W. F. (2002). Asymptotic properties of adaptive designs for clinical trials with delayed response. *Ann Stat*, 30(1), 122–139.

Bauer, P. & Köhne, K. (1994). Evaluation of experiments with adaptive interim analyses. *Biometrics*, 50(4), 1029–1041.

Biswas, A. & Angers, J. F. (2020). Discussion on "Predictively consistent prior effective sample sizes," by Beat Neuenschwander, Sebastian Weber, Heinz Schmidli, and Anthony O'Hagan. *Biometrics*, 76(2), 591–594.

Biswas, A., Bandyopadhyay, U., & Bhattacharya, R. (2007). Response-Adaptive Designs in Phase III Clinical Trials. In *Statistical Advances in the Biomedical Sciences* Chapter 3, (pp. 33–53). John Wiley and Sons, Ltd.

Biswas, A. & Bhattacharya, R. (2016). Response-adaptive designs for continuous treatment responses in phase iii clinical trials: A review. *Stat Methods Med Res*, 25(1), 81–100.

Biswas, A., Bhattacharya, R., & Park, E. (2016). On a class of optimal covariate-adjusted response adaptive designs for survival outcomes. *Stat Methods Med Res*, 25(6), 2444–2456.

Biswas, A. & Mandal, S. (2004). Optimal adaptive designs in phase III clinical trials for continuous responses with covariates. In *mODa 7—Advances in Model-Oriented Design and Analysis: Proceedings of the 7 th International Workshop on Model-Oriented Design and Analysis held in Heeze* (pp. 51–59).: Springer.

Bose, M. & Biswas, A. (2023). Sample sizes required to estimate the protective efficacy of a vaccine when there is an unequal allocation of individuals across the vaccine and placebo groups. *Stat Methods Med Res*, 32(10), 1859–1879.

Brittain, E. & Schlesselman, J. J. (1982). Optimal allocation for the comparison of proportions. *Biometrics*, 38(4), 1003–1009.

Carpenter, T. E. (2001). Use of sample size for estimating efficacy of a vaccine against an infectious disease. *Am J Vet Res*, 62(10), 1582 – 1584.

Chan, I. S. F. & Bohidar, N. R. (1998). Exact power and sample size for vaccine efficacy studies. *Commun Stat - Theory Methods*, 27(6), 1305–1322.

Chen, M. H. & Shao, Q. M. (1999). Monte Carlo estimation of Bayesian credible and HPD intervals. *J Comput Graph Stat*, 8(1), 69–92.

Cheng, T. C. & Biswas, A. (2008). Maximum trimmed likelihood estimator for multivariate mixed continuous and categorical data. *Comput Stat Data Anal*, 52(4), 2042–2065.

Cox, D. R. (1972). Regression models and life-tables. *J R Stat Soc Ser B Methodol*, 34(2), 187–202.

Cox, D. R. (1975). Partial likelihood. *Biometrika*, 62(2), 269–276.

Das, S. & Bhattacharya, R. (2022). An optimal multiarmed response adaptive design for survival outcome with independent censoring. *Biom J*, 64(1), 165–185.

Dupont, W. D. & Plummer, W. D. (1990). Power and sample size calculations: A review and computer program. *Control Clin Trials*, 11(2), 116–128.

Durham, L. K., Longini, Ira M., J., Halloran, M. E., Clemens, J. D., Azhar, N., & Rao, M. (1998). Estimation of vaccine efficacy in the presence of waning: Application to cholera vaccines. *Am J Epidemiol*, 147(10), 948–959.

Eisele, J. R. (1994). The doubly adaptive biased coin design for sequential clinical trials. *J Stat Plann Inf*, 38(2), 249–261.

Eisele, J. R. & Woodroofe, M. B. (1995). Central limit theorems for doubly adaptive biased coin designs. *Ann Stat*, 23(1), 234–254.

Ella, R., Reddy, S., Jogdand, H., et al. (2021). Safety and immunogenicity of an inactivated SARS-CoV-2 vaccine, BBV152: interim results from a double-blind, randomised, multicentre, phase 2 trial, and 3-month follow-up of a double-blind, randomised phase I trial. *Lancet Infect Dis*, 21, 950–961.

Fan, T. H. & Berger, J. O. (1992). Behaviour of the posterior distribution and inferences for a normal mean with t prior distributions. *Stat Risk Model*, 10(1-2), 99–120.

Feikin, D. R., Higdon, M. M., Abu-Raddad, L. J., et al. (2022). Duration of effectiveness of vaccines against SARS-CoV-2 infection and COVID-19 disease: results of a systematic review and meta-regression. *Lancet*, 399(10328), 924–944.

Ferdinands, J. M., Rao, S., Dixon, B. E., et al. (2022). Waning of vaccine effectiveness against moderate and severe covid-19 among adults in the us from the VISION network: test negative, case-control study. *BMJ*, 379.

Gail, M. (1974). Power computations for designing comparative Poisson trials. *Biometrics*, 30(2), 231–237.

Genest, C. & Zidek, J. V. (1986). Combining probability distributions: A critique and an annotated bibliography. *Stat Sci*, 1(1), 114–135.

Ghazy, R. M., Ashmawy, R., Hamdy, N. A., et al. (2022). Efficacy and Effectiveness of SARS-CoV-2 Vaccines: A Systematic Review and Meta-Analysis. *Vaccines*, 10(3).

Ghosh, J., Delampady, M., & Samanta, T. (2006). Large sample methods. In *An Introduction to Bayesian Analysis* Chapter 4. New York: Springer.

Greenwood, M. & Yule, G. U. (1915). The Statistics of Anti-Typhoid and Anti-Cholera Inoculations, and the Interpretation of Such Statistics in General. *Proc R Soc Med*, 8(Sect Epidemiol State Med), 113–194.

Halabi, S. & Singh, B. (2004). Sample size determination for comparing several survival curves with unequal allocations. *Stat Med*, 23(11), 1793–1815.

Halloran, M. E., Haber, M., & Longini Jr, I. M. (1992). Interpretation and estimation of vaccine efficacy under heterogeneity. *Am J Epidemiol*, 136(3), 328–343.

Harrell, F. & Lindsell, C. (2020). Statistical design and analysis plan for sequential parallel-group RCT for COVID-19. *[Internet]* https://hbiostat.org/proj/covid19/bayesplan.

Hsieh, F. & Lavori, P. W. (2000). Sample-size calculations for the Cox proportional hazards regression model with nonbinary covariates. *Control Clin Trials*, 21(6), 552–560.

Hu, F. & Rosenberger, W. F. (2003). Optimality, Variability, Power: Evaluating Response-Adaptive Randomization Procedures for Treatment Comparisons. *J Am Stat Assoc*, 98(463), 671–678.

Hu, F., Zhang, L., Cheung, S. H., & Chan, W. S. (2008). Doubly adaptive biased coin designs with delayed responses. *Can J Stat*, 36(4), 541–559.

Hu, F. & Zhang, L. X. (2004). Asymptotic Properties of Doubly Adaptive Biased Coin Designs for Multitreatment Clinical Trials. *Ann Stat*, 32(1), 268–301.

Hu, F., Zhang, L. X., & He, X. (2009). Efficient randomized-adaptive designs. *Ann Stat*, 37(5A), 2543–2560.

Ivanova, A. (2003). A play-the-winner-type urn design with reduced variability. *Metrika*, 58, 1–13.

Ivanova, A. & Rosenberger, W. F. (2001). Adaptive designs for clinical trials with highly successful treatments. *Drug Inf J*, 35(4), 1087–1093.

Jennison, C. & Turnbull, B. W. (2015). Adaptive sample size modification in clinical trials: start small then ask for more? *Stat Med*, 34(29), 3793–3810.

Joseph, L., M'Lan, C. E., & Wolfson, D. B. (2008). Bayesian sample size determination for binomial proportions. *Bayesian Anal*, 3(2), 269–296.

Joseph, L., Wolfson, D. B., & Du Berger, R. (1995). Sample size calculations for binomial proportions via highest posterior density intervals. *J R Stat Soc Ser D Stat*, 44(2), 143–154.

Katz, D., Baptista, J., Azen, S. P., & Pike, M. C. (1978). Obtaining Confidence Intervals for the Risk Ratio in Cohort Studies. *Biometrics*, 34(3), 469–474.

Koopman, P. A. R. (1984). Confidence Intervals for the Ratio of Two Binomial Proportions. *Biometrics*, 40(2), 513–517.

Lachin, J. M. & Foulkes, M. A. (1986). Evaluation of sample size and power for analyses of survival with allowance for nonuniform patient entry, losses to follow-up, noncompliance, and stratification. *Biometrics*, 42(3), 507–519.

Lal, H., Cunningham, A. L., Godeaux, O., et al. (2015). Efficacy of an adjuvanted herpes zoster subunit vaccine in older adults. *N Engl J Med*, 372(22), 2087–2096.

Malec, D. (2001). A closer look at combining data among a small number of binomial experiments. *Stat Med*, 20(12), 1811–1824.

Marshall, R. J. (1988). Bayesian analysis of case-control studies. *Stat Med*, 7(12), 1223–1230.

Melfi, V. & Page, C. (1998). Variablility in Adaptive Designs for Estimation of Success Probabilities. *Lect Notes-Monogr Ser*, 34, 106–114.

Melfi, V. F., Page, C., & Geraldes, M. (2001). An Adaptive Randomized Design with Application to Estimation. *Can J Stat*, 29(1), 107–116.

M'Lan, C. E. (2002). *Bayesian sample size calculations for cohort and case-control studies*. Doctoral dissertation. McGill University.

M'Lan, C. E., Joseph, L., & Wolfson, D. B. (2006). Bayesian sample size determination for case-control studies. *J Am Stat Assoc*, 101(474), 760–772.

Morita, S., Thall, P. F., & Müller, P. (2008). Determining the Effective Sample Size of a Parametric Prior. *Biometrics*, 64(2), 595–602.

Neuenschwander, B., Capkun-Niggli, G., Branson, M., & Spiegelhalter, D. J. (2010). Summarizing historical information on controls in clinical trials. *Clin Trials*, 7(1), 5–18.

Neuenschwander, B., Roychoudhury, S., & and, H. S. (2016). On the use of Co-Data in Clinical Trials. *Stat Biopharm Res*, 8(3), 345–354.

Neuenschwander, B., Weber, S., Schmidli, H., & O'Hagan, A. (2020). Predictively consistent prior effective sample sizes. *Biometrics*, 76(2), 578–587.

O'Neill, R. T. (1984). Sample sizes for estimation of the Odds Ratio in unmatched case-control studies. *Am J Epidemiol*, 120(1), 145–153.

O'Neill, R. T. (1988). On sample sizes to estimate the protective efficacy of a vaccine. *Statistics in Medicine*, 7(12), 1279–1288.

Orenstein, W. A., Bernier, R. H., Dondero, T. J., Hinman, A. R., Marks, J. S., Bart, K. J., & Sirotkin, B. (1985). Field evaluation of vaccine efficacy. *Bull World Health Organ*, 63(6), 1055–1068.

Orenstein, W. A., Bernier, R. H., & Hinman, A. R. (1988). Assessing Vaccine Efficacy in the field: Further observations. *Epidemiol Rev*, 10(1), 212–241.

Pennello, G. & Thompson, L. (2007). Experience with Reviewing Bayesian Medical Device Trials. *J Biopharm Stat*, 18(1), 81–115.

Pham-Gia, T. (2000). Distributions of the ratios of independent beta variables and applications. *Commun Stat - Theory Methods*, 29(12), 2693–2715.

Polack, F. P., Thomas, S. J., Kitchin, N., et al. (2020). Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *N Engl J Med*, 383(27), 2603–2615.

Proschan, M. A. & Hunsberger, S. A. (1995). Designed extension of studies based on conditional power. *Biometrics*, 51(4), 1315–1324.

Qin, L., Gilbert, P. B., Corey, L., McElrath, M. J., & Self, S. G. (2007). A Framework for Assessing Immunological Correlates of Protection in Vaccine Trials. *J Infect Dis*, 196(9), 1304–1312.

Rosenberger, W. F. & Hu, F. (2004). Maximizing power and minimizing treatment failures in clinical trials. *Clin Trials*, 1(2), 141–147.

Rosenberger, W. F., Stallard, N., Ivanova, A., Harper, C. N., & Ricks, M. L. (2001). Optimal Adaptive Designs for Binary Response Trials. *Biometrics*, 57(3), 909–913.

Savage, L. J. (1971). Elicitation of Personal Probabilities and Expectations. *J Am Stat Assoc*, 66(336), 783–801.

Schoenfeld, D. (1981). The asymptotic properties of nonparametric tests for comparing survival distributions. *Biometrika*, 68(1), 316–319.

Schoenfeld, D. A. (1983). Sample-size formula for the proportional-hazards regression model. *Biometrics*, 39(2), 499–503.

Schoenfeld, D. A. & Richter, J. R. (1982). Nomograms for calculating the number of patients needed for a clinical trial with survival as an endpoint. *Biometrics*, 38(1), 163–170.

Spiegelhalter, D. J., Abrams, K. R., & Myles, J. P. (2003). Prior Distributions. In *Bayesian Approaches to Clinical Trials and Health-Care Evaluation* Chapter 5, (pp. 139–180). John Wiley and Sons, Ltd.

Thall, P., Fox, P., & Wathen, J. (2015). Statistical controversies in clinical research: scientific and ethical problems with adaptive randomization in comparative clinical trials. *Ann Oncol*, 26(8), 1621–1628.

Tsiatis, A. A. (1981). A large sample study of Cox's regression model. *Ann Stat*, 9(1), 93–108.

Van der Vaart, A. W. (1998). Bayes procedures. In *Asymptotic Statistics* (pp. 138–152). Cambridge University Press.

Voysey, M., Clemens, S. A. C., Madhi, S. A., et al. (2021). Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *Lancet*, 397(10269), 99–111.

Walsh, E. E., Jr, R. W. F., Falsey, A. R., et al. (2020). Safety and immunogenicity of two RNA-based Covid-19 vaccine candidates. *N Engl J Med*, 383(25), 2439–2450.

Zhang, D. & Quan, H. (2009). Power and sample size calculation for log-rank test with a time lag in treatment effect. *Stat Med*, 28(5), 864–879.

Zhang, L. & Rosenberger, W. F. (2006). Response-Adaptive Randomization for Clinical Trials with Continuous Outcomes. *Biometrics*, 62(2), 562–569.

Zhang, L. & Rosenberger, W. F. (2007). Response-adaptive randomization for survival trials: the parametric approach. *J R Stat Soc Ser C Appl Stat*, 56(2), 153–165.

Zhang, L. X., Hu, F., & Cheung, S. H. (2006). Asymptotic Theorems of Sequential Estimation-Adjusted Urn Models. *Ann Appl Probab*, 16(1), 340–369.