TECHNIQUES FOR COMPARING EFFICACY AND COST-EFFECTIVENESS OF CANCER THERAPIES, AND IMPROVED INFERENCE TOOLS

A Dissertation
Presented to
The Academic Faculty

by

Mina Plamenova Georgieva

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Industrial and Systems Engineering

Georgia Institute of Technology

August 2018

TECHNIQUES FOR COMPARING EFFICACY AND COST-EFFECTIVENESS OF CANCER THERAPIES, AND IMPROVED INFERENCE TOOLS

Approved by:

Professor Brani Vidakovic, Advisor
School of Industrial and Systems Engineering
Georgia Institute of Technology

Professor David Goldsman
School of Industrial and Systems Engineering
Georgia Institute of Technology

Professor Benjamin Haaland, Advisor
School of Medicine
University of Utah

Dr. Mirjana Brockett
School of Biological Sciences
Georgia Institute of Technology

Professor Yao Xie
School of Industrial and Systems Engineering
Georgia Institute of Technology

Date Approved: April 30, 2018

To my family.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my advisors, Professors Ben Haaland and Brani Vidakovic, for their guidance and immense support. I am thankful to you for nurturing curiosity and passion for research in me. I certainly would have not been here without your devoted encouragement and patience. You have not only been my academic advisors, but also incredible mentors and role models to me. It has been a privilege to work alongside you.

I also want to thank Professors Mirjana Milosevic-Brockett, Yao Xie, and David Goldsman for generously offering their time to serve as committee members. Your support and kindness throughout my doctoral studies means a lot to me. Special thanks to my other collaborators, Joao Paulo da Silveira Nogueira Lima (from Institute of Cancer Research), Gilberto de Lima Lopes (from University of Miami), Akhil Chopra and Chong Ming Yeo (from Johns Hopkins Singapore International Medical Centre), and Pedro Aguiar Junior (from Federal University of Sao Paulo); it has been a pleasure to work with all of you.

I would like to thank the many wonderful faculty and staff members in ISyE for giving me the continued opportunity and support to succeed at all phases of my doctoral program. In particular, Professors Santanu Dey, Yajun Mei, Alejandro Toriello, Craig Tovey, and Roshan Vengazhiyil; thank you for the wonderful classes I had the opportunity to take from you. I greatly admire your dedication and am incredibly grateful for the impact you have had on my life. I also want to thank Gamze Tokol-Goldsman, who I had the absolute pleasure to work with for several semesters as a teaching assistant. Thank you for

your support and kindness. Special thanks to Amanda Ford for making my life at ISyE so much easier.

I would like to thank the many friends and colleagues that I have made during my four years as a Ph.D. student at Georgia Tech. In particular, I want to thank Timur Tankayev, Seyma Guven-Kocak, Tony Yaacoub and Luke Marshall, who shared time, knowledge and laughter with me over the last four years, and made my time at Georgia Tech so wonderful and memorable. There are countless other wonderful people in ISyE who made this experience worthwhile; I will be forever grateful for our time together.

Last, but definitely not least, my appreciation and gratitude goes to my parents and my sister for their continuous love and support throughout my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

AA	Abiraterone Acetate
ADT	Androgen Deprivation Therapy
AR	Androgen Receptor
BRAFi	BRAF inhibitors
BRAFi-MEKi	dual BRAF-MEK inhibitors
CARB+PAC	Carboplatin plus Paclitaxel
CARB+PEM	Carboplatin plus Pemetrexed
CIS+GEM	Cisplatin plus Gemcitabine
CrI	Credible Interval
CRPC	Castration-Resistant Prostate Cancer
CTLA-4i	CTLA-4 blockade
CTLA-4i-PD-1i	dual CTLA-4-PD-1 inhibitor
DP	Dirichlet Process
EoL	End-Of-Life
FDA	Food and Drug Administration
ICERs	Incremental Cost-Effectiveness Ratios
KM	Kaplan-Meier
MCMC	Markov chain Monte Carlo
MEKi	MEK inhibitors
MLE	Maximum Likelihood Estimator
MME	Method of Moments Estimator
NCT	National Clinical Trial

NICE	National Institute of Clinical Excellence
NSCLC	Non-Small Cell Lung Cancer
OS	Overall Survival
PD-1i	PD-1 blockade
PEM Maint	Pemetrexed maintenance
PFS	progression-free survival
PrI	Predictive Interval
QALYs	Quality-Adjusted Life-Years
RCT	Randomized Control Trial
RR	Response Rate
SEER	Surveillance Epidemiology and End Results

SUMMARY

This thesis focuses on two separate topics, one lying at the intersection of health care and statistics, and the other one rising from classical statistical inference. Chapters 2 through 4 address the first topic. They explore and improve techniques for comparing both efficacy and cost-effectiveness of cancer therapies. Chapter 5 focuses on the second topic. It proposes a new estimator for the number of binomial experiments when the success probability is unknown.

Chapter 2 of my thesis establishes an overall ranking of efficacy of possible interventions in patients with advanced or metastatic melanoma within a Bayesian setting. Currently, chemotherapy is established as the standard of care for melanoma, but is often associated with poor responses and short survival. However, recent groundbreaking discoveries in tumor biology and immune surveillance have yielded effective molecularly targeted therapies and immune agents. These new treatments have changed the therapeutic scenario to a completely new reality of high response rates, prolonged disease control, and the possibility of talking of a cure for some patients. These positive results have opened new avenues in the treatment of melanoma patients and, as expected, added layers of complexity to management of those patients. We perform a network meta-analysis in a hierarchical Bayesian random-effects model to assess the role of immunotherapies and targeted therapies. We also evaluate the impact of immunotherapy biomarkers within a hierarchical Bayesian setting with a view to support and improve the therapeutic decision-making process.

Chapter 3 evaluates indirectly the effectiveness of two treatments for advanced castration-resistant prostate cancer (CRPC). Prostate cancer is the most commonly diagnosed cancer. It eventually progresses to CRPC. CRPC is one of the leading cause of cancer-related deaths among men in developed countries. Two novel androgen receptor pathway inhibitors, abiraterone acetate and enzalutamide, have recently become available. They have been developed with the aim of prolonging survival, minimizing complications, and maintaining or improving quality of life in patients with advanced or metastatic CRPC. However, these two treatment options have not been compared head to head against each other in a prospective randomized fashion. In order to choose the optimal treatment and the optimal sequencing of treatments, we perform two analyses. The first one is a comparative effectiveness study within a Bayesian hierarchical setting. The second one is a sequencing assessment of treatments in the context of exponential survival models, informed by Bayesian meta-analyses with between and within study variance components.

Chapter 4 proposes an improved methodology for conducting both meta-analysis and secondary data analyses based on randomized controlled trials. One of the deficiencies inherent to traditional methodology is the lack of individual patient-level data which serves as a basic ingredient for secondary analyses. This shortcoming is handled by recovering the raw time-to-event data through the inverted Kaplan-Meier equations and simulations. The recovered survival distributions are then modeled within a Bayesian semi-parametric framework. We use a hierarchical Dirichlet Process to model discrete-time event probabilities across the time-line up to last follow-up, and a truncated Weibull model to model the tail of the distribution. This approach avoids assumption about the shape of the survival distributions up to the last follow-up time, allows incorporation of censored data,

and accommodates study-to-study heterogeneity. The parametric nature of the Weibull model on the other hand is well suited to making inferences about the survival curve in the absence of data. Finally, patient-level disease trajectories are modeled using a Bayesian Markov model. We demonstrate this methodology using simulations and a study on advanced non-small cell lung cancer.

Finally, Chapter 5 presents a new approach to the binomial n problem, which concerns the estimation of the number of binomial experiments when the success probability is unknown. Some real-life situations, where the problem arises, include the estimation of the number of unreported crimes as well as the number of undetected software errors. Due to its inherent instability, the problem remains fundamentally difficult. Furthermore, neither one of the two parameters of the binomial distribution are unbiasedly estimable when both are unknown. We present an efficient method of estimating the number of trials using a beta-binomial MLE approach. In the absence of replications, when inference about the parameter of interest is not possible, we present a Bayesian approach applied in the context of contingency tables.

CHAPTER 1. INTRODUCTION

In recent years, many new scientific and technological advancements have emerged, and an overwhelming amount of data has become available. On the scientific and technological front, there has been an exponential progress and advancement in many fields. Some discoveries in the past 15 years include reprogramming of stem cells [1], confirming the existence of dark matter [2], producing self-driving cars [3], and others. On the data front, facilitated by the internet of things, the role of data has rapidly evolved. The astonishing volume and variety of data, that has recently been produced, has transformed the world into so called data-driven reality.

Science, technology and data go hand in hand. Data is a key ingredient to developing new scientific and technological tools. Science and technology are necessary to produce, collect, process and understand data. The high-impact nature of the three fields together has provided many opportunities. Together, they are integral components of the decision making and policy making process. However, sometimes limitations on the scientific and technological front lead to knowledge gaps on the data front. Scientific and technological research is often restricted by funding, time, resources, or all three together. These limitations present a challenge because knowledge gaps can in turn hinder the decision-making process and prevent further scientific and technological progress. This thesis focuses on overcoming the knowledge gap in the context of statistical methods.

1.1 Part I

The first part of this thesis focuses on bridging certain knowledge gaps related to cancer care. The goal of Chapters 2 through 4 is to support and improve the therapeutic decision-making process and direct future scientific and technological cancer research efforts.

In recent years, there has been a wave of dramatic successes in the research and treatment of cancer. This progress has been the direct result of revolutionary scientific and technological advances as well as new development in data analysis. A few such novel therapeutic/ diagnostic development approaches include immunotherapy, tumor-agnostic therapy, adoptive T cell therapy, as well as gene therapy. The pace at which the US Food and Drug Administration (FDA) is approving new cancer treatments is unprecedented. In 2016 alone, the FDA approved five uses for immune checkpoint inhibitors in cancer care [4]. From November 2016 through October 2017, the FDA approved a record eighteen new cancer treatments and thirteen uses of cancer therapies [5]. With an increased availability of treatments, physicians face an increasing number of treatment options. However, usually, only a small fraction of the treatments are directly compared against one another in a prospective randomized fashion. Trials are expensive. Funding for clinical research, on the other hand, is limited. Additionally, clinical trials take years to conduct, and cost oncologists' work and patients' lives. Often, conducting studies is not feasible. This is problematic when deciding between compelling treatments that haven't been directly compared before or when choosing the best treatment for a particular patient when an overall ranking of all possible treatments is lacking.

The primary goal of Chapter 2 is to establish an overall ranking of efficacy of possible interventions in patients with advanced or metastatic melanoma. Malignant melanoma is one of the most aggressive types of cancer [6]. Before recent therapeutic advances, once the disease progressed to a metastatic stage, it was almost always fatal [7, 8]. Recent groundbreaking discoveries in tumor biology and immune surveillance have yielded effective molecularly targeted therapies and immune agents in patients who have reached metastasis. Immunotherapies are treatments that boost the immune system. Immunotherapy was named *Advance of the Year* in ASCO's 2017 cancer progress report [4]. Molecularly targeted therapies target specific disease genes or proteins associated with them. Immunotherapies and targeted therapies have rapidly changed the outlook for cancer patients. They have achieved high response rates, prolonged disease control and improved survival [9-13]. Both strategies have changed the therapeutic scenario of advanced melanoma, turning the clinical decision-making a challenging task. With these major advances in research and multiple options now available, a better understanding of how all available treatments compare to each other is needed for selecting the right treatment for a particular patient. However, only a handful of those treatments have been compared directly against each other in a clinical study setting. To fill this knowledge gap, this chapter presents extended comparisons of immunotherapies and targeted therapies for advanced melanoma by incorporating direct and indirect evidence from sixteen published trials. Additionally, we evaluate the impact of certain expressions and mutational status on immunotherapy efficacy.

Chapter 3 performs a comparative effectiveness analyses between two compelling treatments in advanced castration-resistant prostate cancer (CRPC) that haven't been

directly compared before. Prostate cancer poses a significant health problem today. It is the second most common cancer among men [14]. In 2017, prostate cancer alone accounted for almost one in five new cancer diagnoses [15]. Most prostate cancer patients with metastases eventually progress to CRPC within a median of one year [14]. As a result of research efforts over the past decade, two novel treatments, abiraterone acetate and enzalutamide, have recently emerged for the treatment of CRPC. Both therapies are androgen receptor pathway inhibitors. Androgen receptor has been shown to play an important role in the development and progression of prostate cancer [16]. Abiraterone and enzalutamide have each been shown to prolong survival [17, 18]. However, these two treatments haven't been directly compared before in a randomized study. To bridge the gap between practitioners and patients, we compare indirectly the effectiveness using evidence from four randomized trials. We also investigate the optimal sequence of treatments.

Finally, Chapter 4 assesses the cost-effectiveness of pembrolizumab as compared to chemotherapy in patients with advanced non-small cell lung cancer (NSCLC). Lung cancer is one of the most common causes of cancer related death [15]. It is the single leading cause of cancer death among those aged 40 years or older [15]. Due to relatively late diagnosis of lung cancer cases, scientific advances for lung cancer have been slower in contrast to most other cancers [15]. As of 2013, platinum-based chemotherapy was standard treatment for most patients with NSCLC [19]. However, recent scientific advances have led to the discovery of new paradigms for the treatment of NSCLC. In particular, immunotherapies have shown promising results. One such novel treatment is pembrolizumab. In late 2016, preliminary results of a new randomized trial showed superiority of pembrolizumab over standard platinum-based combination chemotherapy [20]. The FDA granted

pembrolizumab accelerated approval for the treatment of patients with metastatic NSCLC [21]. Pembrolizumab has shown impressive clinical results, but analysis of cost-effectiveness of new therapies is imperative to ensure that they are used in an appropriate and sustainable manner. Currently, such cost-effectiveness analysis is lacking. We assess the cost-effectiveness of pembrolizumab in both the United States, and the United Kingdom.

Throughout chapters 3 and 4, we address one additional issue, that is the lack of raw time-to-event data at the individual level. Currently, the standard practice in reporting of results from randomized controlled trials is to publish summary statistics, and not the raw data. Alsheikh-Ali et al found that only 9% of original research papers published in high-impact journals made the raw research data available [22]. The summary statistics that are usually reported include efficacy measures such as hazard ratios and odds ratios. And yet, these measures do not constitute sufficient statistics for conducting secondary analysis, such as treatment efficacy analysis, cost-effectiveness analysis, analysis of sequencing of treatments, and others. We overcome this shortcoming by reconstructing the patient-level survival data.

In conclusion, part I of this thesis aims to provide and improve evidence-based knowledge with a view to help practitioners with the development of new policies and practices related to cancer care.

1.2 Part II

The second part of this thesis bridges the knowledge gap in certain scenarios where data cannot be observed due to limitations, and has to be inferred instead. Consider the

following three scenarios. The first one is presented by Draper and Guttman and involves an appliance repair company [23]. The company is interested in estimating the number of a certain type of appliance in use in a certain service area based on the weekly total number of defective appliances sent in for repair. The second scenario is concerned with the estimation of the total number of crimes when many of them remain unreported [24]. Finally, the third one is related to software systems [25]. Often, there are errors introduced in the software development process. Reviewers can inspect for errors, but often a few errors remain undetected, and the estimation of the number of undetected errors becomes an important task. In these three scenarios, we observe partial counts (number of defective appliances, number of reported crimes, and number of detected errors), while the real interest lies in the total unobserved counts (total number of appliances, total number of crimes, and total number of errors respectively). And yet, the probabilities of observing a certain number of defective appliances, or crimes, or errors are also unknown. In Chapter 5 of this thesis, we focus on bridging the gap between what is observed and what cannot be observed, and estimate the number of binomial experiments when the success probability is unknown.

CHAPTER 2. A SYSTEMATIC REVIEW AND NETWORK META-ANALYSIS OF IMMUNOTHERAPY AND TARGETED THERAPY FOR ADVANCED MELANOMA

2.1 Introduction

Recent groundbreaking discoveries in tumor biology and immune surveillance have yielded effective molecularly targeted therapies and immune agents, changing the scenario from one of poor responses and short survival to a completely new reality of high response rates, prolonged disease control, and the possibility of talking of a cure for some patients [9-13]. Blocking the BRAF-MEK pathway–commonly hyperactive in melanoma–has proved worthwhile. A sizeable number of trials have shown that BRAF inhibitors (BRAFi) and MEK inhibitors (MEKi) improve clinical outcomes when compared to chemotherapy [26-32]. The role of the immune system in controlling melanoma is well established and immune checkpoint inhibitors have shown promise in reinvigorating the immune system, successfully showcasing the enormous potential of drugs that manipulate immune surveillance for the first time in oncology [33-37].

These positive results have opened new avenues in the treatment of melanoma patients and, as expected, added layers of complexity to management of patients with advanced disease. A number of studies have compared competing treatments to one another, but an overall ranking of possible interventions is lacking. The number of options has grown markedly and defining the best therapeutic plan for a particular patient is now a formidable task. This Bayesian network meta-analysis of randomized controlled trials aims

to establish relative efficacy of immunotherapy, molecularly targeted therapies, and chemotherapy, either alone or in combination, in patients with advanced or metastatic melanoma with a view to support and improve the therapeutic decision-making process.

2.2 Patients and Methods

2.2.1 Search strategy

We searched PubMed, Embase, Clinicaltrials.gov, Cochrane Central Register of Controlled Trials, World Health Organization International Trial Registry, drugs at FDA, and Society of Melanoma Research, ASCO, ESMO, and ECCO meetings using a combination of broad terms related to melanoma and drug therapy, namely melan*, random*, immunotherapy, BRAF*, MEK*, and chemotherapy (full list of terms in Appendix A). References in recovered studies and relevant reviews were also screened. Databases were searched from their inception until December 21st 2015. No language restrictions were applied. We followed a predefined protocol (PROSPERO number CRD42016038618) in accordance with the PRISMA guideline for network meta-analysis [38].

2.2.2 Study selection

We searched databases and assessed eligibility of studies based on abstracts and full texts, resolving disagreements by consensus. Eligible studies were (1) randomized controlled trials enrolling patients with metastatic or advanced melanoma and describing outcomes of interest, (2) randomized patients to chemotherapy, targeted therapy against the BRAF/MEK axis or immunotherapy (not vaccine, viral therapy or biochemotherapy),

and (3) BRAF and/or MEK inhibitor trial restricted inclusion to patients known to harbor BRAF mutations. Second-line BRAF-MEK inhibitor studies were eligible if the first-line therapy had not been BRAF-targeted therapy. Studies with insufficient follow-up (≤6 months) or comparing different chemotherapy regimens were excluded. In the case of duplicated publication on the same study, the most up-to- date data were used. We acknowledged that inclusion criterion (4) would exclude NRAS-mutated patients.

2.2.3 Data extraction

We retrieved data from randomized control trial (RCT) full publications and relevant appendices guided by an extraction form. The items of interest were: trial name, first author, year of publication, number of patients, length of follow-up, methodology details (randomization, allocation concealment and blinding methods, use of intention to treat analysis), intervention details (drugs, doses, length of use), patient characteristics (median age, performance status, previous therapy, if any) and outcomes of interest (overall survival, progression-free survival, response rate). Disagreements were resolved by consensus.

2.2.4 Outcomes of interest

Hazard ratios for overall survival (OS) and progression-free survival (PFS), and odds ratios for response rate (RR), were collected or calculated for all included RCTs. We abstracted data from original intention-to-treat multivariate analysis whenever possible; thus, avoiding those derived from landmark analysis or solely based on median comparisons. We adhered to the definition of progression and the criteria used by each trial [39].

2.2.5 Data synthesis and statistical analysis

The comparison of treatments was performed by incorporating both direct and indirect effects within a Bayesian network meta-analysis. Standard meta-analysis is a method of combining evidence from multiple trials of a single comparison into a single effect size. The key here is that traditional meta-analysis does not allow the comparison of treatments if they have not been previously compared directly in a RCT.

The idea behind network meta-analysis was only recently proposed and generalized by Bucher (1997) and Hasselblad (1998) [40, 41]. The term *network meta-analysis* was later coined in 2002 by Lumley, who proposed the application of linear mixed model in the presence of multiple treatments [42]. The network meta-analysis model allows the assessment of relative effectiveness of two treatments when they have not been compared directly in a RCT but have each been compared to other treatments. It strengthens inference regarding relative efficacy of treatments by synthesizing both direct and indirect evidence into a single effect size. Additionally, it facilitates the simultaneous ranking of all treatments and provides a global estimate of comparative treatment effectiveness. Most importantly, it allows the estimation of within- and between-study heterogeneity and the detection of inconsistency between randomized trials.

The method proposed by Lumley is restricted to trials with only two arms. To overcome this limitation, Lu and Ades extended the meta-analysis model developed by Smith, Spiegelhalter and Thomas [43], and proposed a hierarchical Bayesian network meta-analysis for multi-arm studies based on Markov Chain Monte Carlo algorithm [44].

We performed network meta-analysis within a hierarchical Bayesian random-effects model, with relative efficacy measures, hazard and odds ratios, analyzed on the log-scale and random effects for study:

$$\log \theta_{i,j} \sim \text{Normal}\left(x'_{i,j}\beta, \sigma_i^2 + \tau^2\right)$$

where

$$\theta_{i,j}: \text{ hazard ratio (odds ratio) } j \text{ reported in study } i$$

$$x_{i,j}: \text{ treatment contrast } j \text{ in study } i$$

$$\beta: \text{ vector of treatment effects relative to chemotherapy}$$

$$\sigma_i^2: \text{ within} - \text{study variance for study } i$$

$$\tau^2: \text{ between} - \text{study variance in treatment comparisons.}$$

The distribution of all parameters was weighted by a distribution of prior beliefs. Parameters were given either non- or weakly informative priors letting the pooled data dominate the posterior distribution. Weakly informative priors were used for the mean treatment effects, placing 95% of the prior probability on hazard (odds) ratios between 1/10 (1/20) and 10 (20), so that the pooled data dominated the posterior distribution. In particular, the effectiveness of treatment k relative to chemotherapy β_k was given the following weakly informative prior:

$$\beta_k \sim \begin{cases} \mathrm{N}\left(0, \left(\frac{\log(10)}{2}\right)^2\right) & \text{for hazard ratios} \\ \mathrm{N}\left(0, \left(\frac{\log(20)}{2}\right)^2\right) & \text{for odds ratios} \end{cases}.$$

Priors for individual within study variances σ_i^2 were specified via inverse gamma distribution with reported value as its mean and variance proportional to D_i, the number of events for OS or PFS outcomes

$$\sigma_i^2 \sim Inverse\ Gamma\left(\frac{D_i}{2}, \frac{D_i}{2}\hat{\sigma}_i^2\right),$$

where $\hat{\sigma}_i^2$ is the reported variance for study i. For studies that did not report number of events (death for OS and progression or death for PFS), number of events were estimated by proxies as follows: for OS the assumption was that 50% of the randomized patients died, and for PFS 75% had PFS events by study cut-off.

Similarly, between-study variances τ^2 were assigned a weakly informative uniform distribution,

$$\tau \sim \begin{cases} U\left(0, \left(\frac{\log(2)}{2}\right)^2\right) & \text{for hazard ratios} \\ U\left(0, \left(\frac{\log(5)}{2}\right)^2\right) & \text{for odds ratios} \end{cases},$$

which allows hazard (odds) ratios to vary by up to two-fold (five-fold) across studies.

Finally, within-study correlation among the two relative efficacy measures in the three arm trials was modeled as bivariate normal whose marginal distributions matching those described above and having a correlation coefficient, ρ. A non-informative prior distribution

$$\rho \sim U(0,0.95)$$

was taken for ρ.

Estimates from three-arm studies were modeled in the context of a bivariate normal distribution with the same weakly informative prior on the between study variance along with an uninformative prior on the within-study correlation.

Samples from the posterior distribution of the parameters were generated via Markov Chain Monte Carlo implemented through JAGS within R http://mcmc-jags.sourceforge.net/[45-47]. Ten chains were used with the first 100,000 iterations of each discarded as "burn-in". Results are based on 500,000 iterations from each chain, thinned at a lag of 100.

We calculated posterior mean hazard and odds ratios for relative efficacy of each therapy, along with credible 95% intervals, predictive 95% intervals, and probabilities of each treatment being better than a reference were calculated. Therapies which achieved the combined benchmarks (a) overall survival (OS) posterior mean HR ≤ 0.8 with probability better ≥ 80% as compared to chemotherapy, (b) progression-free survival (PFS) posterior mean HR ≤ 0.6 with probability better than chemotherapy ≥ 90%, and (c) response rate (RR) posterior mean OR ≥ 3.0 with probability better than chemotherapy ≥ 95% were deemed to have a *meaningful benefit* as compared to chemotherapy [48].

Additionally, we performed a traditional pairwise meta-analysis for all treatments that have been directly compared in a trial before. We used a model similar to the approach of DerSimonian and Laird published in 1986 [49].

We tested the hypothesis that BRAF mutation status alters relative efficacy of immunotherapy. Interactions between BRAF mutation status and relative efficacies were incorporated in the model. We also tested the hypothesis that PD-L1 expression affects relative efficacy of immunotherapies CTLA-4-PD-1 dual blockage, PD-1 blockage and CTLA-4 blockage. We adhered to the trial definition of PD-L1 positivity.

Study-to-study heterogeneity was summarized using predictive intervals, which provide an interval in which the relevant comparative efficacy measure would be expected to fall for a new study. Ranking and probabilities were calculated based on predicted relative effects drawn from the posterior. Quality of studies was assessed via Cochrane Collaboration's tool for assessing risk of bias in randomized trials [50]. Publication bias was graphically assessed via funnel plot.

2.3 Results

2.3.1 Systematic review

A total of 1750 published or presented titles and abstracts were screened. After duplicated review and discussion, 18 trials on 10 types of therapy, comprising 7596 patients, had their data extracted. All trials were multicentric and reported in English. A sizeable number of trials used chemotherapy (dacarbazine, paclitaxel or temozolomide) as control arm. Trials assessing BRAF-MEK dual blockade used BRAFi as control arm and restricted enrollment to patients harboring BRAF mutations. When dealing with trials comparing MEK-chemotherapy versus chemotherapy, we restricted the data to BRAF-mutated patients. No trial performed a head-to-head comparison of immunotherapy versus

BRAFi. The majority of excluded randomized trials failed to use BRAFi or immunotherapy as active comparator.

Two trials have been omitted from the main analysis as they have not produced relevant data; one comparing dacarbazine to dacarbazine and ipilimumab and other comparing ipilimumab to ipilimumab and sargramostim (available upon request)[36, 51]. Hence, the main analysis gathered data from 16 trials with eight therapeutic nodes and 6849 patients [26-28, 30, 33-35, 52-66].

All included evidence was intention-to-treat, based on standard analyses, from studies with low risk of bias, according to the Cochrane risk of bias tool (provided in Figure 23 in Appendix A). No sign of publication bias was found using the funnel plot (provided in Figure 24 in Appendix A). The schematic flowchart of systematic review is presented in Figure 1. Table 1 summarizes the trails included in the main analyses.

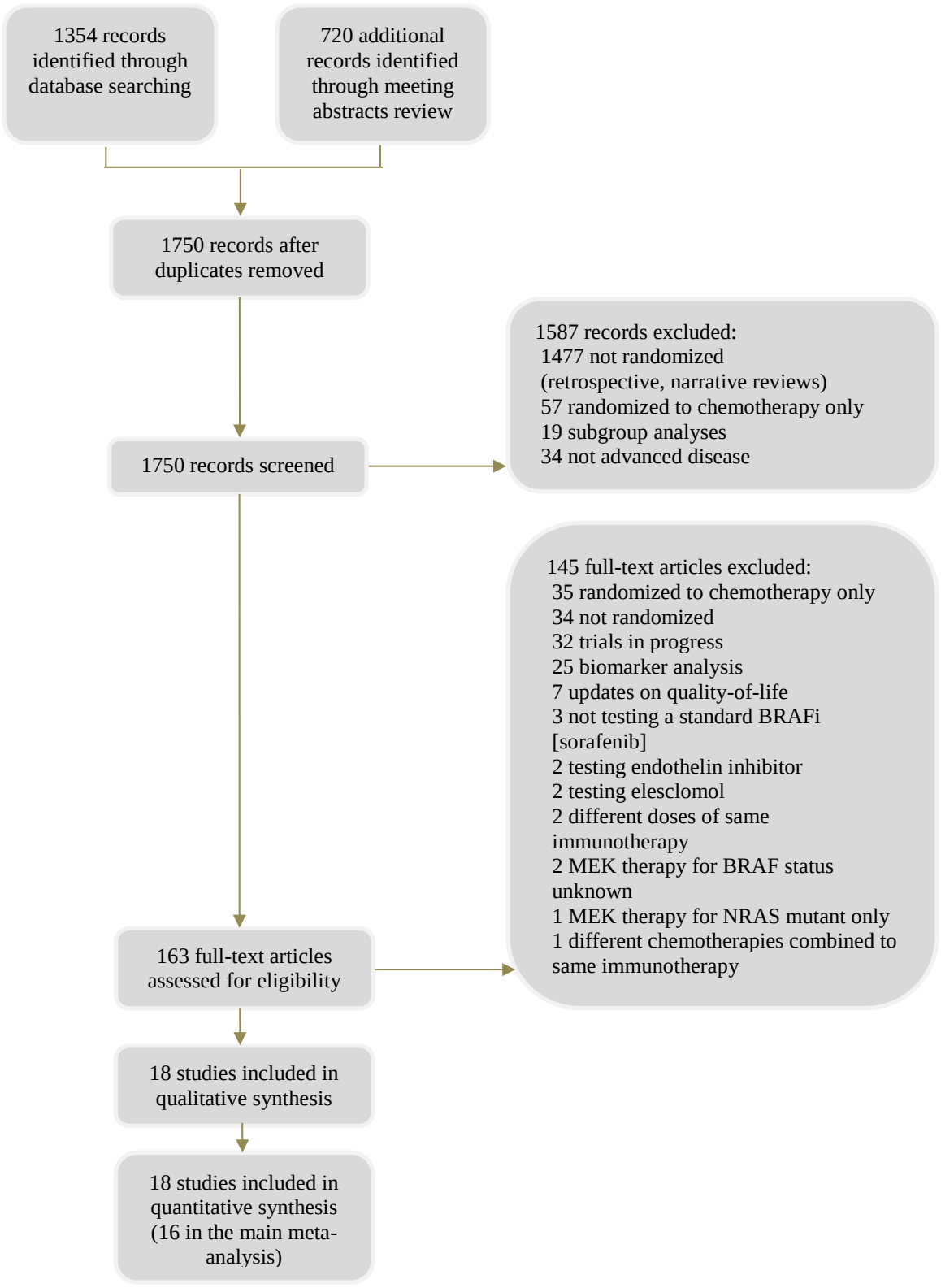

Figure 1. PRISMA flowchart of systematic review of studies included in the Bayesian network meta-analysis

Table 1: Main features of included trials - (A) BRAFi or MEKi trials and (B) Immunotherapy trials.

Study Acronym NCT	Population (line of therapy)	Treatments	N	OS HR (95% CI)	PFS HR (95% CI)	Response (%)
BREAK-3 [28, 54] NCT01227889	Unresectable stage III or IV BRAF V600E mutated (1st or 2nd)	Dabrafenib 150 mg po bd	187	0.77 (0.52-1.13)	0.35 (0.20–0.61)	93 (50)
		DTIC[4]	63	Reference	Reference	4 (6)
BRIM-3 [26, 60] NCT01006980	Previously untreated metastatic BRAF V600E mutated (1st)	Vemurafenib 960 mg po bd	337	0.70 (0.57-0.87)	0.38 (0.32-0.46)	192 (57)
		DTIC[4]	338	Reference	Reference	29 (9)
BRF113220[1] [27, 52] NCT01072175	Metastatic, no previous BRAFi; BRAF mutated (1st, 2nd, 3rd)	Dabrafenib 150 mg po bd + trametinib 2 mg po od	54	0.79 (0.49-1.27)	0.39 (0.25-0.62)	41 (76)
		Dabrafenib 150 mg po bd + trametinib 1 mg po od	54	0.96 (0.57-1.60)	0.56 (0.37-0.87)	27 (50)
		Dabrafenib 150 mg po bd	54	Reference	Reference	29 (54)
coBRIM [56, 58] NCT01689519	Previously untreated; unresectable stage III or IV; BRAF mutated (1st)	Vemurafenib 960 mg po bd + cobimetinib 60 mg po od 3 weeks on 1 week off	247	0.70 (0.55-0.90)	0.58 (0.46-0.72)	172 (70)
		Vemurafenib 960 mg po bd + placebo	248	Reference	Reference	124 (50)
COMBI-d [29, 30] NCT01584648	Previously untreated; unresectable stage IIIC or IV; BRAF mutated (1st)	Dabrafenib 150 mg po bd + trametinib 2 mg po od	211	0.71 (0.55-0.92)	0.67 (0.53-0.84)	144 (68)
		Dabrafenib 150 mg po bd+ placebo po od	212	Reference	Reference	112 (53)
COMBI-v [64, 65] NCT01597908	Previously untreated; metastatic; BRAF mutated (1st)	Dabrafenib 150 mg po bd + trametinib 2 mg po od	352	0.66 (0.53-0.81)	0.61 (0.51-0.73)	226 (64)
		Vemurafenib 960 mg po bd	352	Reference	Reference	180 (51)
METRIC [53, 59] NCT01245062	Unresectable stage III or IV BRAF mutated (no previous BRAFi, MEKi or ipilimumab) (1st or 2nd)	Trametinib 2 mg po od	214	0.54 (0.32-0.92) [no cross-over (0.38; 0.15-0.95)]	0.42 (0.29-0.59)	47 (22)
		DTIC[4] or Paclitaxel5[5]	108	Reference	Reference	9 (8)
NCT00338130 [55]	Previously untreated; unresectable stage III or IV (1st) [2]	Selumetinib 100 mg po bd continuously	45	1.65 (0.91-3.01)	1.24 (00.73-2.10)	5 (11)
		Temozolomide	28	Reference	Reference	3 (11)
NCT00936221 [63]	Previously untreated; advanced BRAF-mutated cutaneous or unknown primary (1st) [2]	Selumetinib 75 po bd + DTIC[4]	45	0.93 (0.57-1.53)	0.63 (0.40-0.98)	18 (40)
		Placebo po bd + DTIC[4]	46	Reference	Reference	12 (26)

Table 1 (Continued)

(B) CheckMate 037 NCT01721746 [33]	Progression after ipilimumab (and BRAFi if BRAF mutated) (2nd or further	Nivolumab 3 mg/kg iv every 2 weeks	272	-	-	38 (32)
		Carbotaxol[6] or DTIC[4]	133	Reference	Reference	5 (11)
CheckMate 066 NCT01721772 [35]	Previously untreated; unresectable, stage III or IV non-uveal, BRAF wild type (1st)	Nivolumab 3 mg/kg iv every 2 weeks + DTIC-placebo	210	0.42 (0.25-0.73)	0.43 (0.34-0.56)	84 (40)
		DTIC[4] + nivo-placebo iv every 2 weeks	208	Reference	Reference	29 (14)
CheckMate 067 NCT01844505 [57]	Previously untreated; unresectable stage III or IV; BRAF mutated (1st)	Nivolumab 3 mg/kg iv every 2 weeks + ipi-placebo iv	316	-	0.57 (0.43-0.76)	138 (44)
		Nivolumab 1 mg/kg + Ipilimumab 3 mg/kg iv both every 3 weeks 4× then Nivolumab 3 mg/kg iv every 2 weeks	314	-	0.42 (0.31-0.57)	181 (58)
		Ipilimumab 3 mg/kg + nivo-placebo iv every 3 weeks 4× then nivo-placebo iv every 2 weeks	315	-	Reference	60 (19)
CheckMate 069 NCT01927419 [34, 67]	Previously untreated; unresectable, stage III or IV (1st)	Nivolumab 1 mg/kg + Ipilimumab 3 mg/kg iv every 3 weeks 4× then Nivolumab 3 mg/kg iv every 2 weeks (BRAF wild type)	72	[3]	0.40 (0.23-0.68)	44 (61)
		Ipilimumab 3 mg/kg + Placebo iv every 3 weeks 4× then Placebo iv every 2 weeks (BRAF wild type)	37	-	Reference	4 (11)
		Nivolumab 1 mg/kg + Ipilimumab 3 mg/kg iv every 3 weeks 4× then Nivolumab 3 mg/kg iv every 2 weeks (BRAF mutated)	23	-	0.38 (0.15-1.00)	12 (52)
		Ipilimumab 3 mg/kg + Placebo iv every 3 weeks 4× then Placebo iv every 2 weeks (BRAF mutated)	10	-	Reference	1 (10)

Table 1 (Continued)

Keynote 002 NCT01704287 [62]	Progression after ipilimumab and BRAFi if BRAF mutated (2nd or 3rd)	Pembrolizumab 2 mg/kg iv every 2 weeks	180	-	0.57 (0.45-0.73)	38 (21)
		Pembrolizumab 10 mg/kg iv every 2 weeks	181	-	0.50 (0.39-0.64)	46 (25)
		DTIC[4] or paclitaxel[b] or temozolomide[7] or carbotaxol[8]	179	Reference	Reference	8 (4)
Keynote-006 NCT01866319 [66]	Unresectable stage III or IV (1st or 2nd)	Pembrolizumab 10 mg/kg iv every 2 weeks	183	0.58 (0.41-0.84)	0.55 (0.42-0.72)	62 (34)
		Pembrolizumab 10 mg/kg iv every 3 weeks	185	0.68 (0.47-0.96)	0.50 (0.38-0.66)	60 (33)
		Ipilimumab 3 mg/kg iv every 3 weeks 4x	181	Reference	Reference	22 (12)
NCT00257205 [61]	Previously untreated; unresectable stage III or IV (1st)	Tremelimumab 10 mg/kg every 90 days	328	0.9 (0.75-1.07)	0.94 (0.81-1.11)	36 (11)
		Temozolomide[7] or DTIC[4]	327	Reference	Reference	32 (10)

NCT, National Clinical Trial (NCT) number found on clinicaltrials.gov; *N*, number of enrolled patients; OS, overall survival; PFS, progression-free survival; HR, hazard ratio; BRAFi, BRAF inhibitor; MEKi, MEK inhibitor; po, oral; od, once a day; bd, twice a day; iv, intravenously; ipi-placebo, placebo matched to ipilimumab; nivo-placebo, placebo matched to nivolumab.
[1]Included patients from randomized part (part C) of the trial.
[2]BRAF mutation-positive data extracted from subgroup analysis.
[3]Data available after systematic review and not included in the meta-analysis.
[4]DTIC: Dacarbazine 1000 mg/kg iv every 3 weeks.
[5]Paclitaxel: Paclitaxel 175 mg/m2 every 3 weeks.
[6]Carbotaxol: Paclitaxel 175 mg/m2 plus carboplatin AUC 5 both iv every 3 weeks.
[7]Temozolomide: temozolomide 200 mg/m2/d 5 days ON every 28 days.
[8]Carbotaxol: Paclitaxel 225 mg/kg plus Carboplatin AUC 6 both iv every 3 weeks.

2.3.2 Quantitative analysis

The 16 trials were grouped across eight therapeutic nodes (6849 patients) according to type of therapy: chemotherapy, CTLA-4 blockade (CTLA-4i), PD-1 blockade (PD-1i), BRAF inhibitors (BRAFi), MEK inhibitors (MEKi), dual BRAF-MEK inhibitors (BRAFi-MEKi), chemotherapy-MEKi, and dual CTLA-4-PD-1 inhibitors (CTLA-4i-PD-1i). Figure 2 describes the network design of treatments' comparison. All standard

chemotherapies (paclitaxel, temozolomide, dacarbazine) were gathered into a single therapeutic node (chemotherapy), with analogous collapse for PD-1 drugs (nivolumab and pembrolizumab). BRAFi and MEKi results are restricted to BRAF mutated patients across all comparisons. The area of the circle in Figure 2 is proportional to the sample size of patients enrolled in each node; the width of connecting lines indicates the number of direct comparisons within trials. The nodes were organized based on the following groupings:

- Chemo: chemotherapy;

- *: MEKi + chemotherapy;

- **: CTLA-4i-GMCSF;

- ***: CTLA-4-chemotherapy;

- Green circles: immunotherapy nodes;

- Orange circles: BRAFi or MEKi-based nodes;

- Blue circle: chemotherapy node.

Number of patients in each node: CTLA-4i: 1172; PD-1i: 1527; CTLA-4i-PD-1i: 409; CTLA-4-chemotherapy: 250; CTLA-4i-GMCSF: 123; MEKi single agent: 259; Chemotherapy: 804; BRAFi single agent: 1390; BRAFi + MEKi: 918; MEKi + chemotherapy: 45. Not all trials described all outcomes (Table 1).

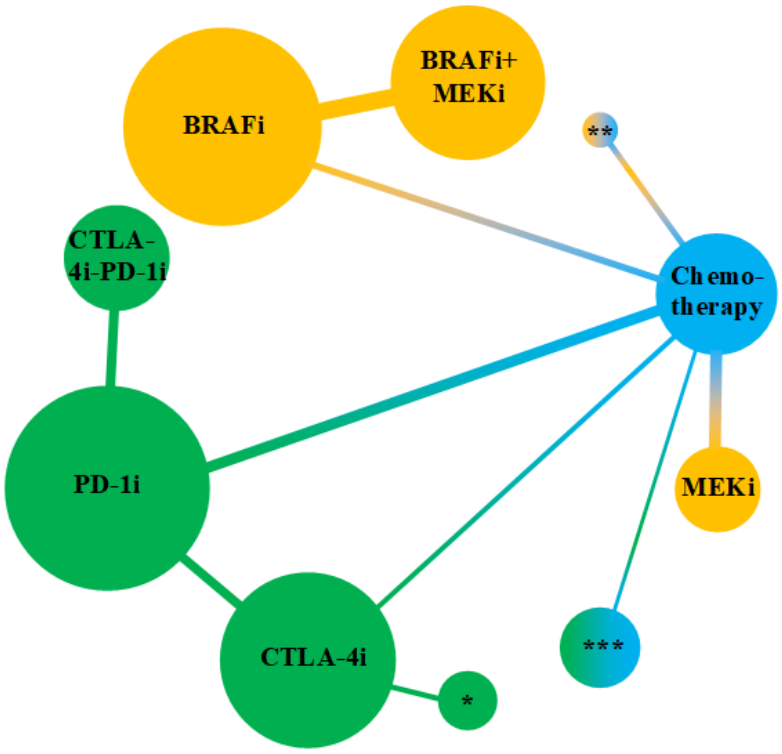

Figure 2. Network diagram of therapeutic nodes.

2.3.3 Efficacy

Three therapies achieved meaningful benefit as compared to chemotherapy: PD-1 blockade, BRAFi-MEKi combination and BRAFi. As evidenced by comparing the prediction and confidence intervals for OS, PFS and RR, study-to-study heterogeneity was present, but broadly had little impact on posterior ranking of treatments.

2.3.4 Overall survival

OS data were available for 12 (of 16) studies including 4817 patients. The results based on traditional pairwise meta-analysis and Bayesian network meta-analysis were aligned with no identifiable signal of inconsistency between indirect and direct approaches.

Three therapies improved OS when compared to chemotherapy, BRAFi-MEKi combination (HR: 0.50; 95% CrI: 0.34–0.74; 95% PrI: 0.31–0.82), PD-1i (HR: 0.52; 95% CrI: 0.36–0.75; 95% PrI: 0.32–0.83), and BRAFi (HR: 0.71; 95% CrI: 0.51–0.97; 95% PrI: 0.46–1.09). PD-1i and BRAFi-MEKi performed similarly (HR: 1.03; 95% CrI: 0.60–1.76; 95% PrI: 0.56–1.90) with probability of BRAFi-MEKi being superior to PD-1i of 55.8%. Both BRAFi-MEKi and PD-1i had high posterior probability of outperforming all competitors. Full comparative OS results are provided in Figure 3. Given high probabilities of outperforming competitor therapies, for PFS and RR, BRAFi-MEKi combination may be optimal for BRAF-mutated patients, whereas PD-1i may be optimal for BRAF wild-type patients or selected BRAF-mutated patients.

Despite the lack of OS data for CTLA-4i-PD-1i combination at the time of systematic review, PFS and RR data were suggestive that CTLA-4i-PD-1i could also achieve meaningful benefit and consequently be a top-ranking option irrespective to BRAF status (see below) [34, 57].

The results based on traditional pairwise meta-analysis were consistent with the results based on the Bayesian network meta-analysis (see Figure 4). Figure 5 displays a comparison of the results of the Bayesian network meta-analysis to the corresponding reported estimates from the 12 randomized studies of melanoma treatments in 4,817 patients.

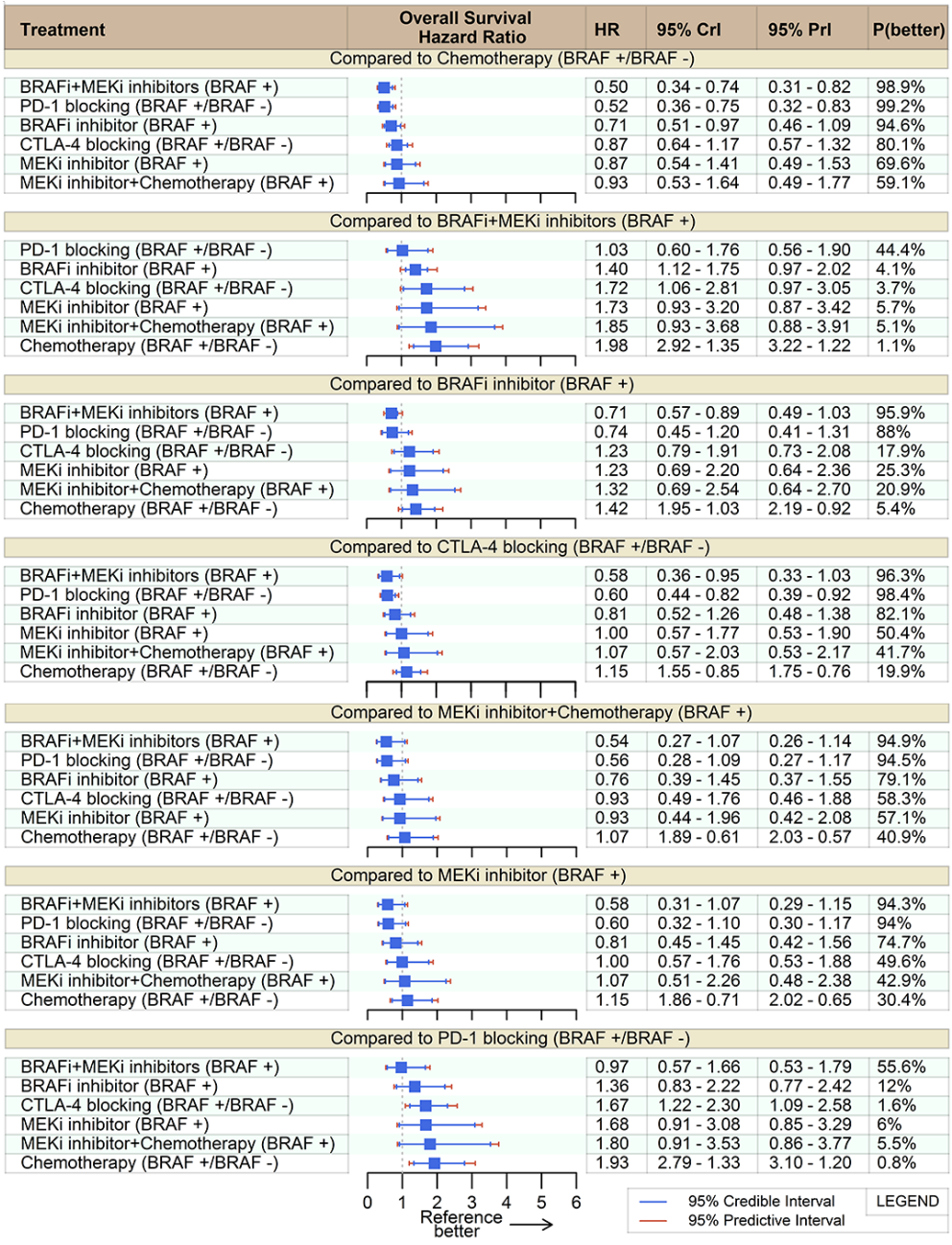

Figure 3. Overall survival network meta-analysis

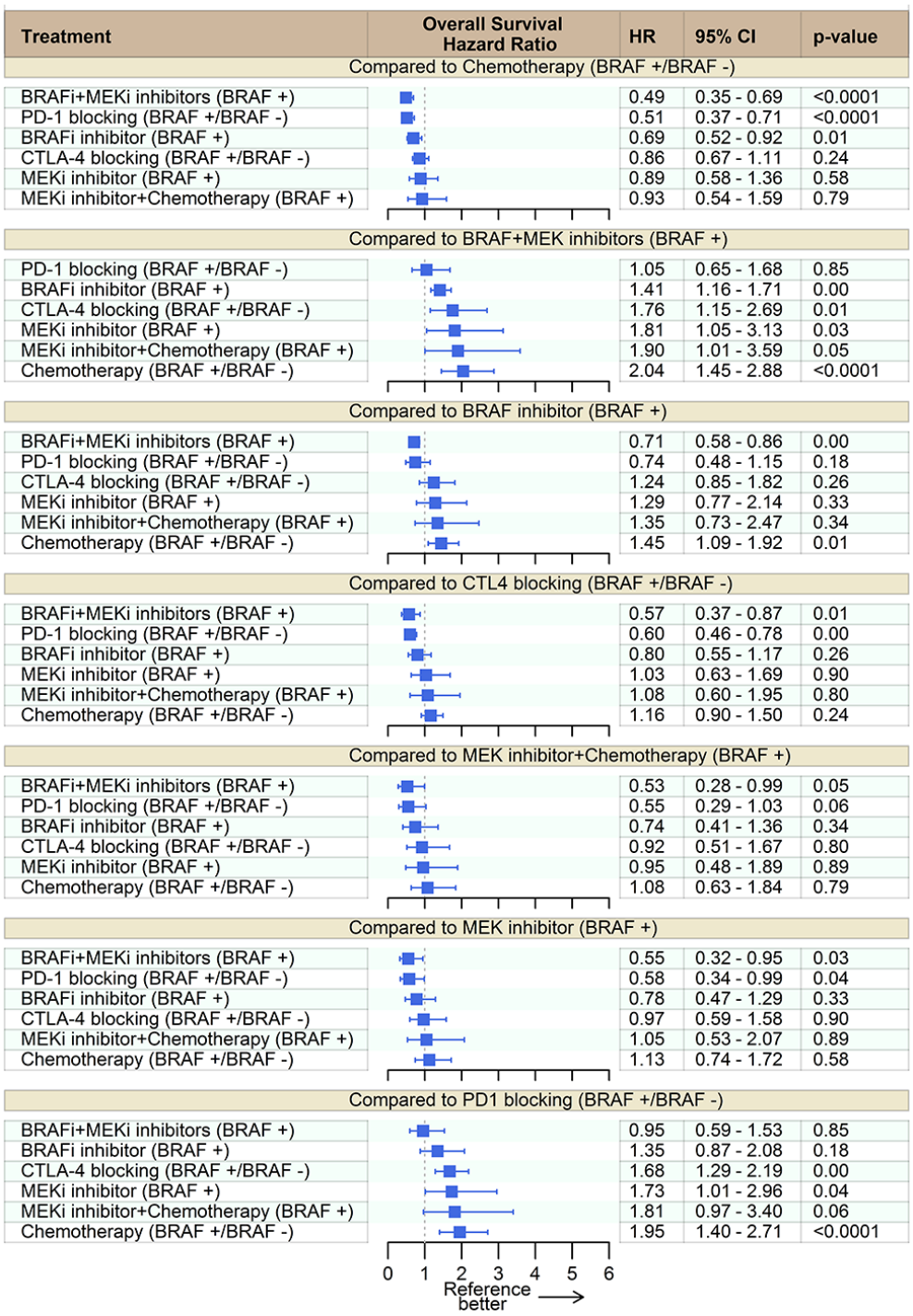

Figure 4. Overall survival traditional meta-analysis

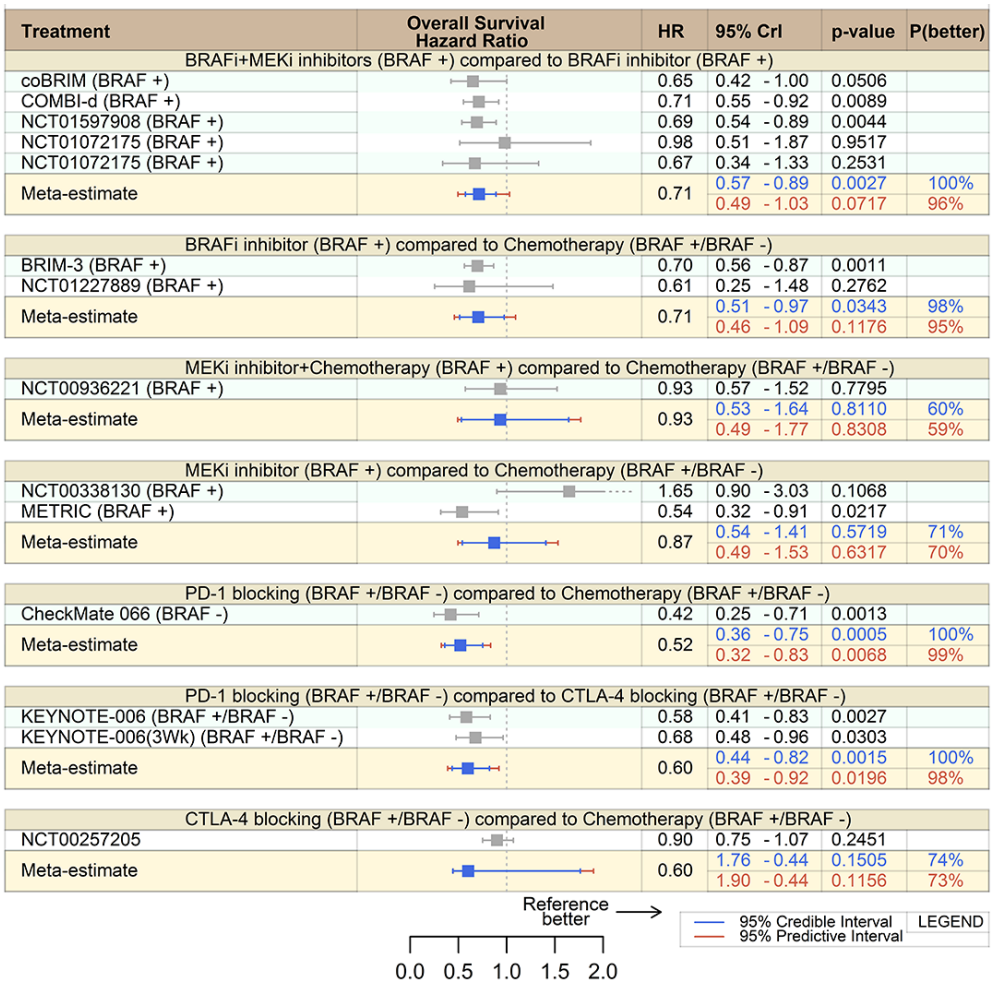

Figure 5. Overall survival details

2.3.5 Progression-free survival

Fifteen trials contributed to the PFS analysis. Worthy of note, the trial comparing tremelimumab (CTLA-4i) to chemotherapy provided 6-month time-restricted PFS data with tumor assessments done at different time points, every 6 weeks in the dacarbazine arm and every 12 weeks in the tremelimumab arm [61]. This study was not included in the PFS analysis.

Four therapies clearly stood better than chemotherapy: BRAFi-MEKi (HR: 0.22; 95% CrI: 0.16–0.31; 95% PrI: 0.14–0.34), CTLA-4i-PD-1i (HR: 0.39; 95% CrI: 0.25–0.6; 95% PrI: 0.23–0.66), BRAFi (HR: 0.39; 95% CrI: 0.29–0.52; 95% PrI: 0.26–0.59), and PD-1i (HR: 0.5; 95% CrI: 0.4–0.64; 95% PrI: 0.34–0.73). Single agent PD-1i and dual CTLA-4i-PD-1i, both outperformed CTLA-4i with corresponding posterior probability of 99.5% (HR: 0.53; CrI: 0.42–0.68) and 99.9% (HR: 0.42; CrI: 0.3–0.57). CTLA-4i had similar performance to chemotherapy (HR: 0.94; CrI: 0.67–1.31).

Dual BRAFi-MEKi yielded the best PFS results with a 96.2% posterior probability of outranking the remaining options, even when compared to CTLA-4i-PD-1i (HR: 0.56; CrI: 0.33–0.97). CTLA-4i-PD-1i and BRAFi stood close as next options (CTLA-4i-PD-1i vs. BRAFi HR: 1.00; 95% CrI: 0.6–1.67), both probably above single agent PD-1i. Full comparative PFS results are provided in Figure 6. Figure 7 contains the results from the traditional meta-analysis. Figure 8 shows that the estimates from the network meta-analysis are consistent with the published studies.

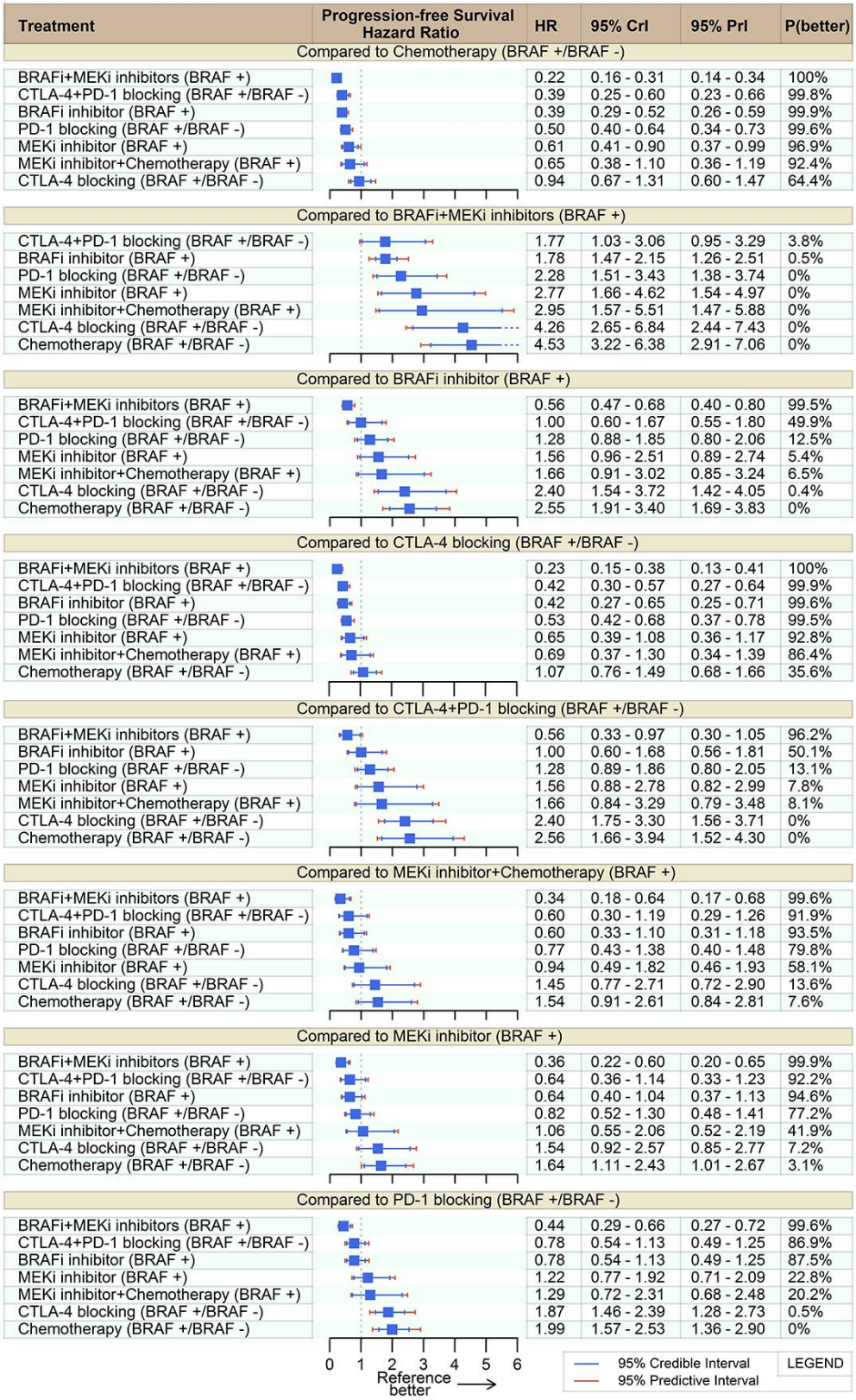

Figure 6. Progression-free survival network meta-analysis

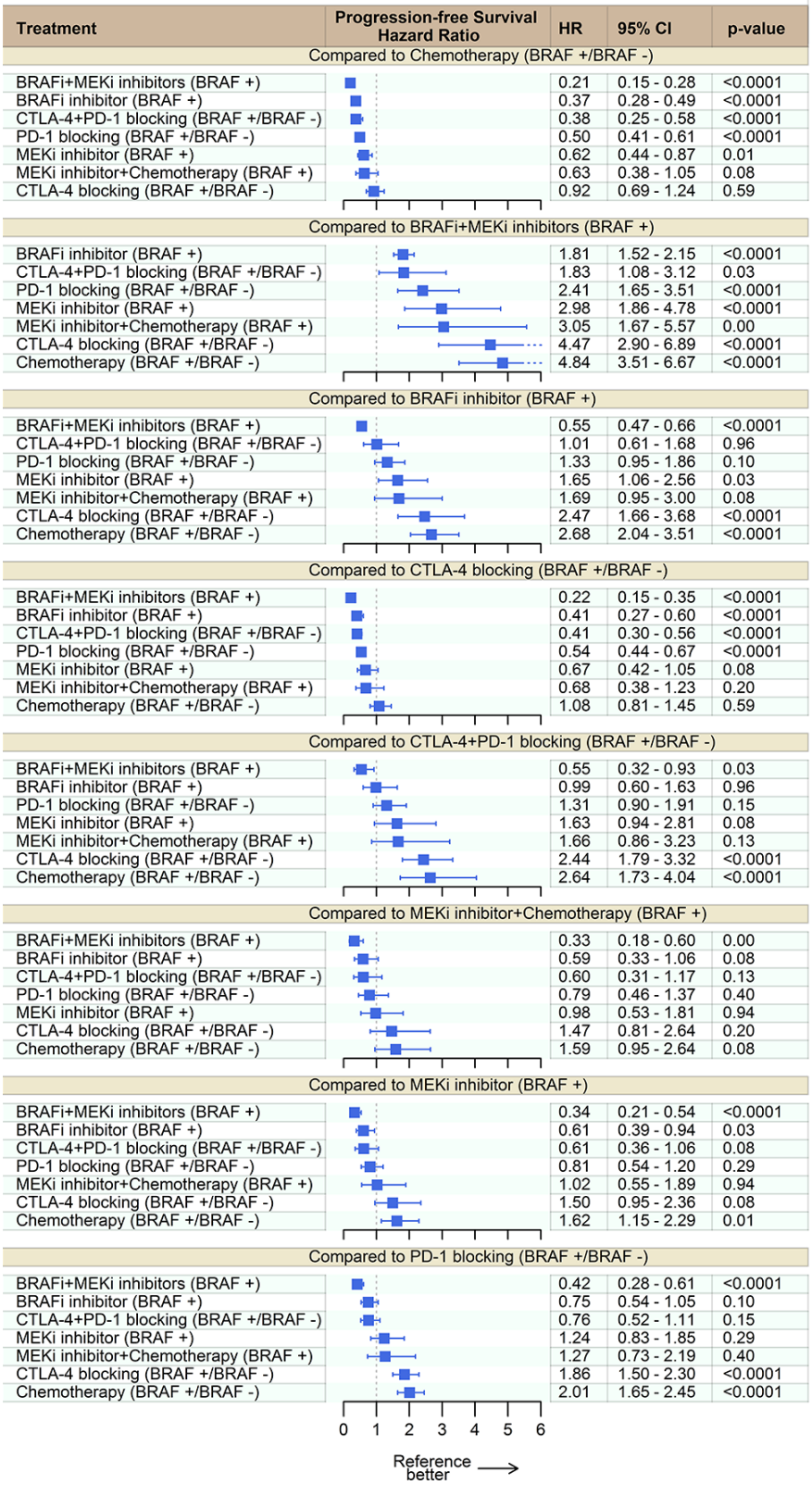

Figure 7. Progression-free survival traditional meta-analysis

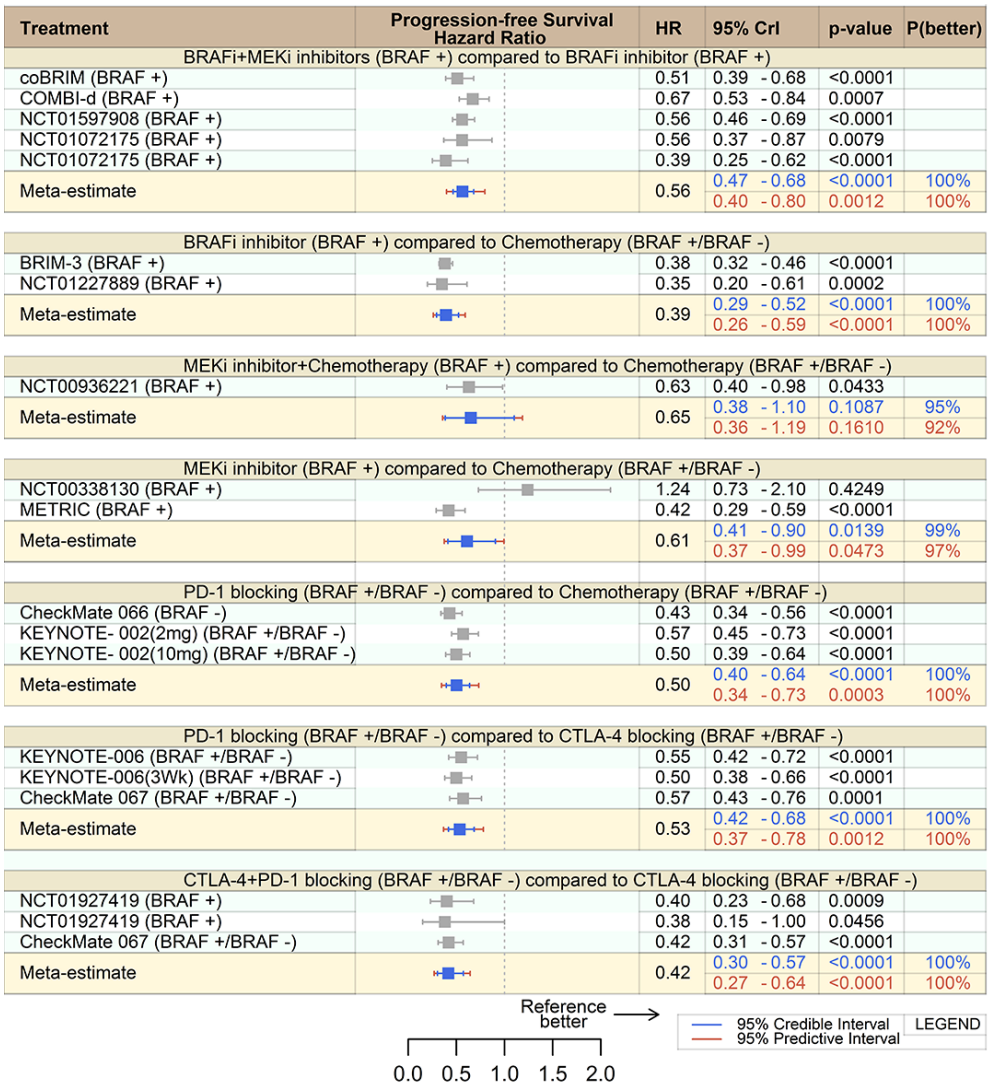

Figure 8. Progression-free survival details

2.3.6 Response rate

RR data were available for all studies. Bearing in mind that response under CTLA-4i can be a late event, we included the tremelimumab versus chemotherapy trial in this analysis. Four therapies led to meaningful benefit (OR ≥ 3.0 and probability better ≥ 95% vs. chemotherapy): BRAFi-MEKi (HR: 19.76; 95% CrI: 10.45–37.35; 95% PrI: 9.19–42.52), BRAFi (HR: 10.78; 95% CrI: 6.24–18.63; 95% PrI: 5.4–21.48), CTLA-4i-PD-1i

(HR: 7.25; 95% CrI: 4.09–12.86; 95% PrI: 3.57–14.7), and PD-1i (HR: 4.32; 95% CrI: 3.07–6.09; 95% PrI: 2.52–7.45). Full comparative RR results are presented in Figure 9. Results from traditional meta-analysis are given in Figure 10, and results from the network meta-analysis are compared to published estimates in Figure 11.

Dual BRAFi-MEKi therapy topped best with at least 97.1% posterior probability of being superior to any other treatment: CTLA-4i-PD-1i (OR: 2.73; CrI: 1.18–6.3), CTLA-4i (OR: 17.2; CrI: 8.31–35.58), PD-1i (OR: 4.57; CrI: 2.24–9.31), MEKi (OR: 8.56; CrI: 3.32–22.04), and BRAFi (OR: 1.83; CrI: 1.37–2.45). For BRAF-mutated patients, the second best option was BRAFi. CTLA-4i-PD-1i dual checkpoint blockade had a 94.3% posterior probability of being superior to single agent PD-1i (OR: 1.68; 95% CrI: 0.99–2.84).

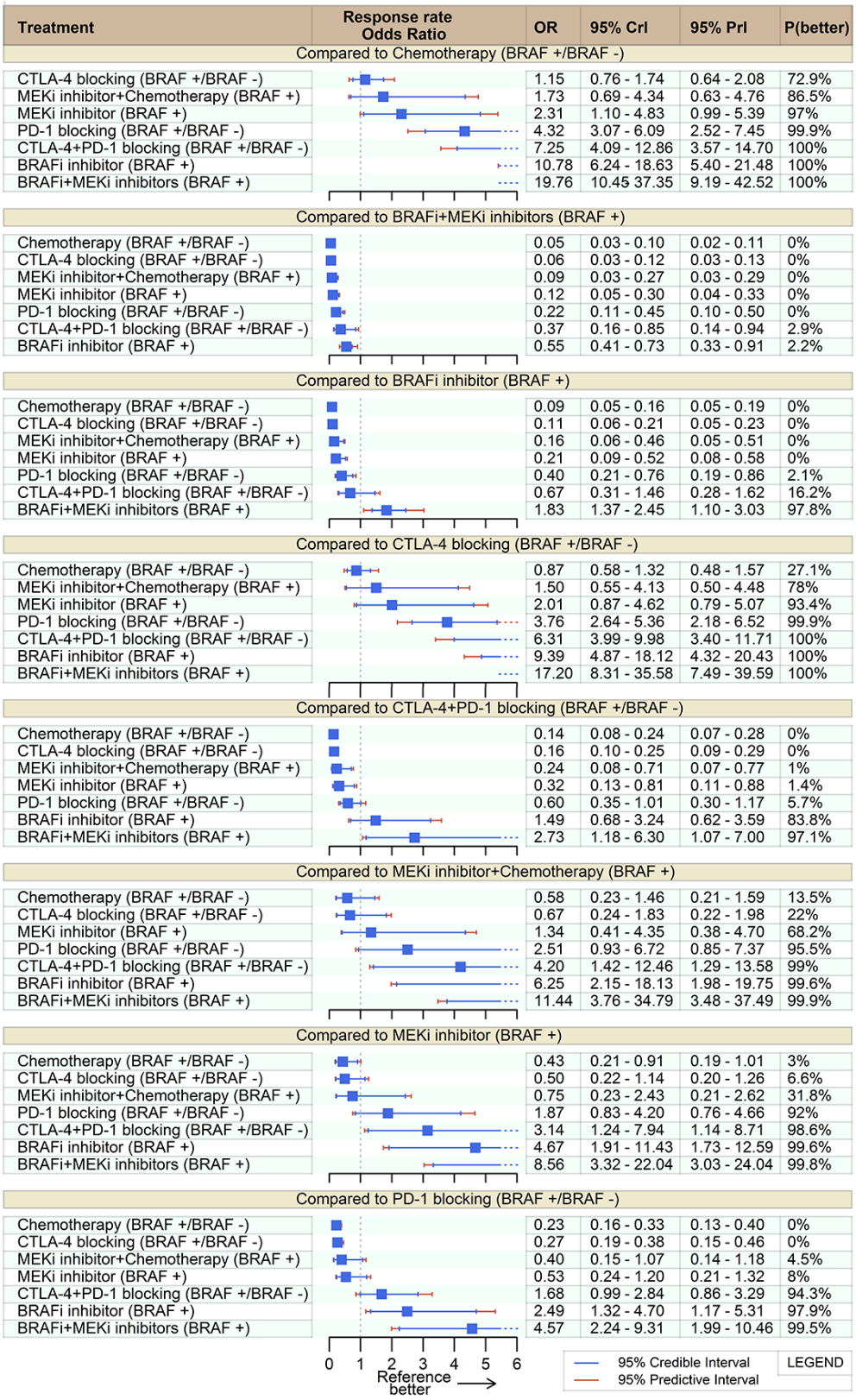

Figure 9. Response rate network meta-analysis

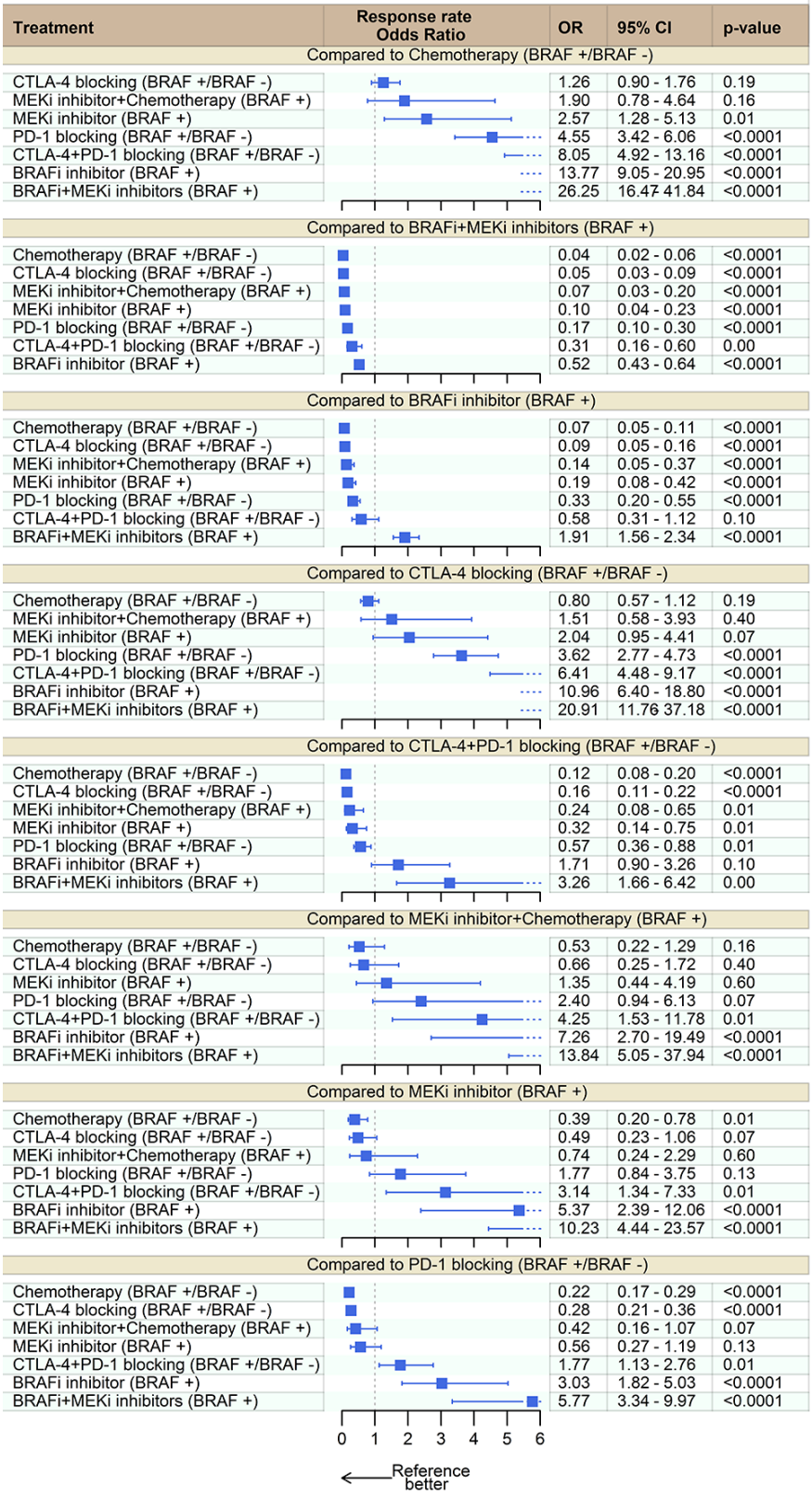

Figure 10. Response rate traditional meta-analysis

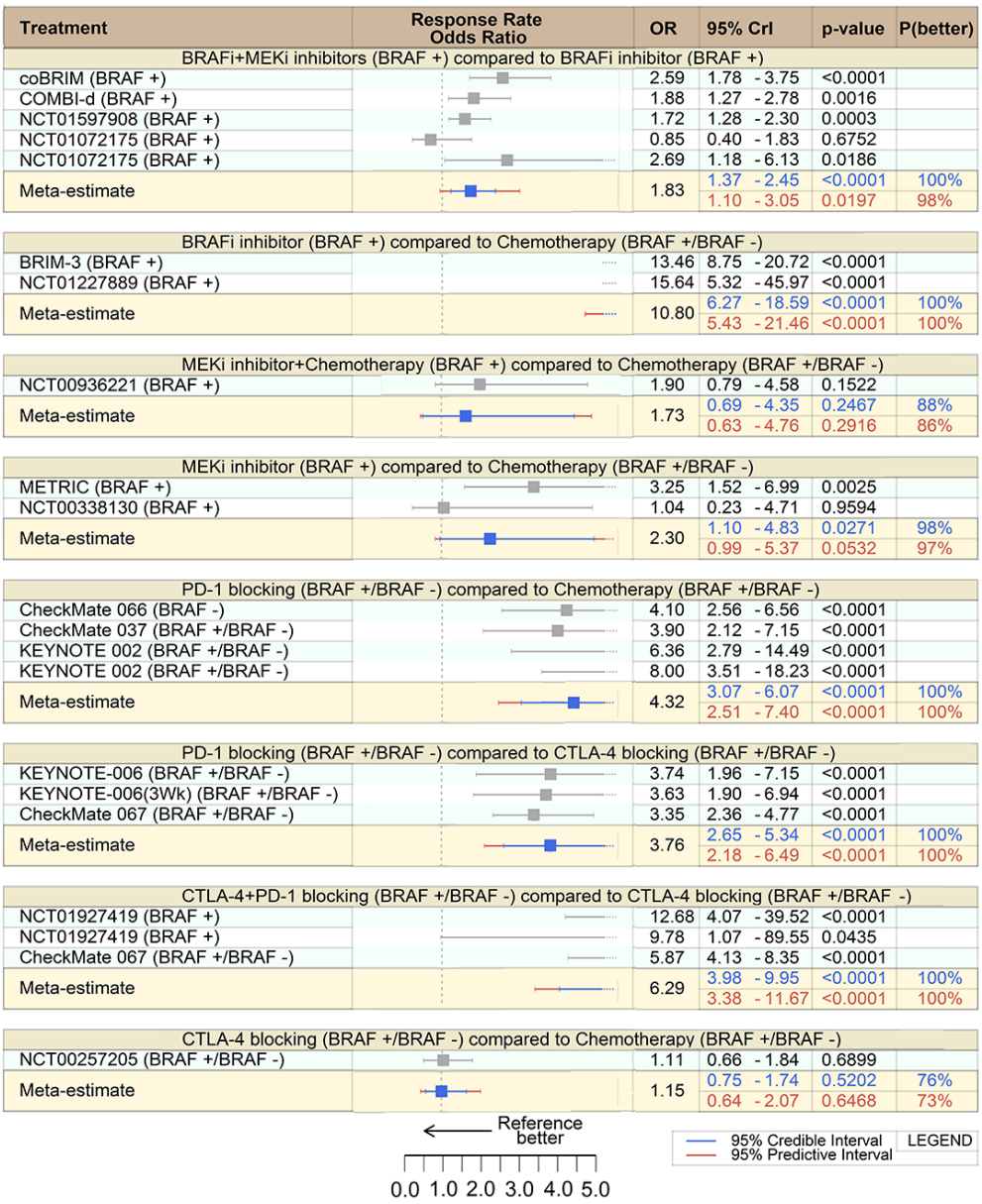

Figure 11. Response rate details

2.3.7 PD-L1 expression and BRAF mutational status as biomarkers of response to immunotherapy

The Bayesian network meta-analysis failed to identify any relevant impact of BRAF mutation status on efficacy of immunotherapy treatments for OS, PFS, or RR in all subsets

sought. The hazard ratios, 95% credible and predictive intervals of BRAF-mutated and wild-type patients were superimposable, which negates any role of BRAF status as a predictor of benefit of immunotherapy (Table 2). The posterior probability that BRAF+ patients had better efficacy of immunotherapies relative to chemotherapy, [P(BRAF+ better)] was from 21% to 50% for OS, 17% to 51% for PFS, and from 16% to 61% for RR. Also, 95% CrIs failed to show any difference according to BRAF mutation status. Results were similar for data selection containing first-line studies with results stratified by BRAF status; credible and predictive intervals did not show evidence of a difference between BRAF wild-type and mutation positive patients in terms of relative efficacy of immunotherapies. As all trials testing BRAFi limited the enrollment of BRAF+ patients, BRAF status was disregarded from the analysis henceforth

Two immunotherapy trials provided information on outcomes according to PD-L1 status [34, 57]. As the definitions of positive and negative tumor PD-L1 expression as well as the laboratory methods used to ascertain them were not homogenous across the two PD-1 trials (Nivolumab: at least 5% of tumor cells with PD-L1 at any intensity at the membrane; Pembrolizumab: >1% tumor cells with membranous PD-L1 expression), we accepted the trials' original cutoffs.

For both PFS and RR, the Bayesian network meta-analysis failed to show any relevant impact of PD-L1 status on efficacy of CTLA-4i-PD-1i, PD-1i, or CTLA-4i. The hazard ratios and 95% CrIs of PD-L1 positive and PD-L1 negative patients overlapped, failing to identify any difference according to PD-L1 status (Table 3). The posterior probability that PD-L1 positive patients had better efficacy under CTLA-4i-PD-1i, PD-1i,

or CTLA-4i, (probability PD-L1+ better) was from 44% to 56% for PFS, and 62% to 83% for RR.

Table 2. Comparison of treatments according to BRAF mutation status

First-line by BRAF mutation												
	Overall Survival				Progression-free Survival				Response rate			
Treatment	HR	95% CrI	95% PrI	P(better)	HR	95% CrI	95% PrI	P(better)	OR	95% CrI	95% PrI	P(better)
CTL-4 blocking	-	-	-	-	1.01	0.12-8.9	0.11-9.07	49%	0.68	0.04-11.05	0.04-12.28	38.97%
CTL-4 + PD-1 blocking	-	-	-	-	0.99	0.11-8.75	0.11-8.93	50.27%	1.46	0.09-23.98	0.08-26.74	60.47%

First-line												
	Overall Survival				Progression-free Survival				Response rate			
Treatment	HR	95% CrI	95% PrI	P(better)	HR	95% CrI	95% PrI	P(better)	OR	95% CrI	95% PrI	P(better)
CTL-4 blocking	0.99	0.05-19.26	0.05-19.47	50.12%	0.54	0.11-2.77	0.1-2.82	77.39%	1.3	0.12-14.34	0.11-15.48	58.83%
CTL-4 + PD-1 blocking	-	-	-	-	0.56	0.1-3.25	0.1-3.3	74.27%	2.18	0.15-31.32	0.14-33.58	71.90%
PD-1 blocking	0.47	0.08-2.98	0.07-3.03	78.84%	0.49	0.12-2.02	0.12-2.07	83.93%	2.8	0.36-21.62	0.33-23.64	83.99%

First- and second-line												
	Overall Survival				Progression-free Survival				Response rate			
Treatment	HR	95% CrI	95% PrI	P(better)	HR	95% CrI	95% PrI	P(better)	OR	95% CrI	95% PrI	P(better)
CTL-4 blocking	0.99	0.05-18.9	0.05-19.21	50.26%	0.59	0.1-3.35	0.1-3.81	72.93%	1.3	0.12-13.99	0.11-15.04	59.00%
CTL-4 + PD-1 blocking	-	-	-	-	0.61	0.1-3.82	0.09-3.95	70.35%	2.2	0.15-31.54	0.14-33.41	71.95%
PD-1 blocking	0.48	0.08-3.04	0.07-3.12	78.38%	0.48	0.11-2.18	0.1-2.29	83.06%	2.79	0.36-21.37	0.34-23.04	84.06%

Table 3: Comparison of treatments according to tumoral PD-1 status

	Progression-free Survival				Response Rate			
Treatment	HR	95% CrI	95% PrI	P(better)	OR	95% CrI	95% PrI	P(better)
CTL-4 blocking	0.84	0.1-6.92	0.1-7.15	56.32%	0.78	0.15-4.08	0.12-5.02	38.46%
CTL-4 + PD-1 blocking	-	-	-	-	0.66	0.11-4.07	0.09-4.89	32.34%
PD-1 blocking	1.19	0.15-9.75	0.14-10.03	43.51%	0.47	0.1-2.21	0.08-2.73	16.69%

2.4 Discussion

This meta-analysis synthesizes the wealth of information on immunotherapy and BRAFi/MEKi for advanced melanoma, producing a ranking of the drugs currently available. The network approach attempts to circumvent the absence of direct comparisons among the many available options, notably the comparison of immunotherapy to BRAF-MEK inhibition and among immunotherapies. The present meta-analysis suggests that dual BRAFi-MEKi is the most effective in improving OS, PFS, and RR of BRAF-mutated patients, outperforming other treatments.

Among the BRAF-MEK axis inhibition options, single-agent BRAFi ranked below BRAFi-MEKi combination, but could still offer higher benefits than single MEKi. These findings may prompt inquiry into how to manage dose reduction of MEKi and BRAFi in the event of toxicities likely to be caused by both drugs. However, clinically relevant this question is, it is beyond the scope of our study to provide such practical guidance.

Appraising the PFS and RR scenarios, it was conceivable that BRAFi-MEKi would dominate them, as BRAF-MEK inhibition was already known to produce frequent and rapid responses, whereas immunotherapy may take longer to produce sustained tumor shrinkage and even lead to unconventional response patterns not properly captured by the standard response assessments [39, 68, 69]. CTLA-4i epitomized the immune response pattern: failed to improve PFS and RR when compared to chemotherapy, but prolonged OS, as the original trials suggested [36, 61]. Our findings underscore the perception that, standard PFS assessment may not be the best way to capture anti-tumor activity of immunotherapy. Nevertheless, it is noteworthy that dual BRAFi-MEKi also stood as the best option with regard to OS, even when compared to single-agent PD-1i.

Notwithstanding the BRAF-MEK inhibition dominance, PD-1 blockade still ranked high in terms of OS, PFS, and RR. Hence, PD-1i may be an attractive option for BRAF wild-type patients and even for BRAF-mutated patients, as it ranked in second to BRAFi-MEKi. OS results for combined CTLA-4-PD-1 immune checkpoint inhibition are not yet mature and longer follow-up may change the order of top-ranked therapies. Some very recent results have started to become available with promising long-term survivorship with dual immune checkpoint blockade [67]. Those findings seem to embody the preliminary reports of prolonged disease control under immunotherapy [13].

We could not confirm the role of PD-L1 as a biomarker of response to PD-1i-based therapy. As currently tested, tumor PD-L1 expression did not better inform the patient selection for PD-1-based therapy, both PD-L1 positive and negative patients derived substantial benefit from PD-1-based therapy. This finding somewhat diverged from the realms of other tumors, showcasing the particular features of immune response within each tumor type [70]. Also, our results failed to show any impact of BRAF status on response to PD-1 therapy, confirming previous findings [71].

Several issues may be implicated on the lack of surrogacy of PD-L1 expression. The simplest one would be statistical power constrained by a small sample size. This indeed could have played a role, however, more than 800 patients—evenly divided between PD-L1 positive and negative— provided data for this analysis. Another possibility would be the use of inadequate cutoffs. To properly assess this, individual patient data would be required. However, even if such data were available, the different antibodies and techniques would require careful consideration. Harmonization of laboratory methods should be enacted first, as is already occurring in lung cancer with the different PD-1/PD-L1 agents.

Lastly, baseline PD-L1 expression at a single tumor site may not be capable of fully capturing the complexity of anti-PD-1-led orchestration of immune system dynamics. It is conceivable that resetting a whole system—in the case of immune system—might be multilayered and continuously changing.

The quest for excellent patient selection is key. Better patient selection transcends optimizing clinical outcomes. It can improve financial resource allocation, a real-world

hurdle to be crossed when new technologies are under consideration. Furthermore, identifying the most likely patients for immunotherapy will spare the nonresponders from fairly toxic therapies. The results of cooperative work on other tumors may enhance our understanding on this important topic [72-76].

Given the number of therapeutic options currently available for advanced melanoma, the sequencing of drugs is another crucial question. The wealth of information organized by this meta-analysis may shed light on the long-term therapeutic plan for melanoma patients. These nuances of clinical management are yet to be defined. However, we believe that clinicians will now be better informed for the decision-making process. Definitive results on sequencing of the various therapeutic options will add to the knowledge base [77, 78].

A major clinical concern is the effectiveness of immunotherapy after progressing under previous BRAF-MEK treatment. Two immunotherapy trials enrolled patients who had progressed while on BRAF-targeted therapy [33, 62]. No sign of loss of efficacy was identified with the use of PD-1 drug among this group of patients as compared to BRAF therapy-naïve patients. Such findings must be further validated and the opposite drug order also appraised, the latter being the question of active trials [78].

This meta-analysis faced several shortcomings inherent to the methodology applied. We had no access to individual patient data, precluding a more detailed appraisal of outcomes and patients' characteristics. This is especially true for assessment of the role of PD-L1 expression, volume of disease, and presence of other known prognostic markers [79, 80]. We concentrated on efficacy foregoing analysis of toxicity, another major

practical concern on clinical grounds. The different cutoffs used for defining PD-L1 status hindered a more robust analysis of its relevance. The absence of overall survival data for CTLA-4-PD-1 trials is a major shortcoming and hopefully more data will become available in the near future [67]. Also, for the sake of simplicity, we analyzed all drugs in the same therapeutic node as identical (for instance tremelimumab and ipilimumab as CTLA-4i prototypes). Furthermore, the duration of response could not be formally assessed as the original trials lacked enough information for a comprehensive appraisal.

Another concern was the publication and trial quality biases. We sought the most relevant databases in order to collect all published and presented trials so far, checked their references and references from relevant reviews and followed Cochrane′s guidelines on the topic. Also, we preplanned the inclusion of BRAFi or immunotherapy trials in order to concentrate on the most promising therapies; hence, some randomized trials testing other targeted therapies, such as sorafenib, oblimersen, or endothelin inhibitors were not meta-analyzed. Trials enrolling personalized therapy to other targets, such as NRAS-mutant tumor, were not included [81]. With regard to the quality of trials included, nearly all trials were ascribed as high quality according to the Cochrane risk of bias tool, with the lack of placebo as the commonest source of likely bias.

Furthermore, it is conceivable that gathering different drugs with different doses and regimens in the same node could lead to heterogeneity, and some heterogeneity was found among the several comparisons made. Nevertheless— and most importantly—direct comparison results were in line with the network results and the impact of heterogeneity on the ranking of therapy options was minimal.

In spite of all those shortcomings listed above, we were able to formally compare different therapies and provide a clear rank of efficacy of the many available options for advanced melanoma. Abstracting all this sizeable amount of information, combined BRAFi-MEKi-targeted therapy seems to be a sound option at the present—even in light of emerging results of immune therapy—for BRAF-mutant patients. Longer follow-up in dual immune checkpoint trials coupled with further analysis of immune markers have the potential to further enhance outcomes in advanced melanoma.

CHAPTER 3. ABIRATERONE OR ENZALUTAMIDE IN CASTRATION-RESISTANT PROSTATE CANCER: INDIRECT COMPARISON

3.1 Introduction

Prostate cancer is the most commonly diagnosed cancer and the second leading cause of cancer-related deaths among men in developed countries [82]. In the United States, according to the National Cancer Institute's Surveillance Epidemiology and End Results (SEER) database, prostate cancer prevalence in 2012 was estimated at approximately 890 per 100,000 men. Approximately 14.0% of men will be diagnosed with prostate cancer at some point during their lifetime.

A significant majority of prostate cancers are diagnosed at an early-localized stage; however, some patients will relapse with disseminated disease while others are diagnosed with advanced cancer at initial presentation [83].

Prostate cancer cells are dependent on androgen receptor (AR) signaling for growth and survival. Androgen-deprivation therapy is the standard of care for advanced or metastatic prostate cancer, and has been for decades [84-87]. Even though more than 90% of prostate cancer patients initially respond to androgen deprivation therapy; many tumors become refractory and castration-resistant with time.

Multiple active treatment modalities have been developed for men with advanced castration-resistant prostate cancer (CRPC) with the aim of prolonging survival,

minimizing complications, and maintaining or improving quality of life. These agents have distinct modes of action and include chemotherapeutic agents, such as docetaxel and cabazitaxel, the immunotherapeutic agent, sipuleucel-T, the bone targeting alpha-emitting radionuclide, radium-223 chloride, as well as the novel androgen receptor (AR) pathway inhibitors abiraterone acetate (abiraterone) and enzalutamide [88-95].

The development of novel anti-androgens and androgen synthesis inhibitors as a result of research efforts over the past decade show that CRPC remains dependent on AR function for growth by evolving multiple mechanisms to activate receptor signaling such as ligand independent activation of AR, verexpression of the AR receptor, gain of function mutations in AR, and upregulation of androgen biosynthesis enzymes [96].

Abiraterone acetate (AA), a pro-drug of abiraterone, is a selective irreversible inhibitor of the products of the CYP17 gene (including both 17,20-lyase and 17- alpha-hydroxylase), and thereby blocks synthesis of androgens in tumor as well as in the testis and adrenal glands. Enzalutamide is an orally administered, potent next-generation antiandrogen agent that acts at multiple sites in the androgen receptor signaling pathway, including blocking binding of androgen to the androgen receptor, inhibition of nuclear translocation of the androgen receptor, and inhibition of the association of the androgen receptor with nuclear DNA. Unlike abiraterone, concurrent treatment with steroids is not required.

AA and enzalutamide have both been investigated and shown to prolong overall survival in large phase III trials in both the pre- and post-docetaxel settings [91-94]. Large-scale, prospective randomized trials testing the optimal sequencing of these treatments

have not yet been reported, nor have the two agents been compared head to head against each other in a prospective randomized fashion. There is some evidence from a number of small retrospective cohort studies suggesting limited activity of these agents when used in a sequential fashion either before or after docetaxel chemotherapy in advanced CRPC [97-107]. Development of predictive biomarkers to facilitate the selection of patients for a specific therapy or sequence of therapies is the focus of ongoing efforts. Recently, the AR-V7 splice variant, a truncated isoform of the AR that lacks the binding domain of both enzalutamide and AA, was shown to be associated with resistance to both agents as evidenced by inferior PSA50 response rates, PFS and OS [108]. AA and enzalutamide differ in the use of prednisone and in the incidence of toxicities, which can be used in the decision-making process either for upfront and sequential therapy.

To assist practicing clinicians in decision-making, we performed two analyses, a comparative effectiveness study using available evidence from *randomized* studies and a sequencing assessment using additional available evidence from *observational* studies of enzalutamide and AA in a post-AR pathway inhibitor setting. Indirect meta-analyses are often used to provide preliminary guidance when head to head evidence is not available.

In the comparative effectiveness study, enzalutamide and AA were compared indirectly in terms of OS, radiographic PFS, time to PSA progression, PSA response rates (RR), and adverse events in both the pre-docetaxel and post-docetaxel setting [91-94]. In the sequencing assessment, the treatment ordering of enzalutamide and AA are compared in terms of OS in the post-docetaxel setting, using both randomized and observational evidence [92, 93, 97-101, 103, 104, 106].

3.2 Methods

3.2.1 Comparative Effectiveness

Indirect meta-estimates were generated in the context of a Bayesian hierarchical model with study specific efficacy estimates meta-analyzed on the log (of hazard or odds ratio) scale similarly to the approach found in section 2.2.5. The primary endpoint, OS, and secondary endpoints, radiographic PFS, time until PSA progression, and PSA response rate, were modeled as jointly multivariate Gaussian with mean depending on each study's treatment contrast and variance-covariance matrix composed as the sum of a diagonal matrix of within study variances and a full dimensional between study variance-covariance matrix. Treatment effects considered were enzalutamide relative to placebo in the pre-docetaxel setting, AA relative to placebo in the pre-docetaxel setting, modification of both enzalutamide and AA effects in the post-docetaxel setting, and modification of AA effects due to the addition of prednisone in the placebo arm in both the pre- and post-docetaxel settings. A Bayesian perspective is appropriate from a decision making (choosing the best treatment in a particular context) point of view, and allows seamless incorporation of sources of uncertainty. The primary measure of efficacy was posterior probability enzalutamide outperforms AA with prednisone in terms of OS on average. Secondary measures of efficacy were posterior probability enzalutamide outperforms AA with prednisone in terms of OS in an individual study setting, as well as hazard and odds ratios along with 95% credible and predictive intervals. Both pre- and post-docetaxel settings were of interest.

Throughout, priors were selected to reflect the range of realistically plausible parameter values. In particular, priors for enzalutamide and AA relative to placebo in the pre-docetaxel setting placed 95% of their mass on hazard ratios between 1/10 and 10 (odds ratios between 1/400 and 400), priors for the modification of both enzalutamide and AA effect in the post-docetaxel setting placed 95% of their mass on hazard ratios between 1/1.25 and 1.25 (odds ratios between 1/1.5 and 1.5), and priors for the modification of AA effect due to the addition of prednisone in the placebo arm placed 95% of their mass on hazard ratios between 1/1.1 and 1.1 (odds ratios between 1/1.25 and 1.25). Priors for within study variances were taken as inverse gamma with mean equal to the reported (or recalculated) standard errors and variance proportional (conservatively) to each study's total number of deaths. Priors for between study variances were taken as uniform on (0, 0.175), allowing high prior probability of up to twofold differences in hazard and odds ratios across individual study settings, and priors for correlations between endpoints were taken as uniform on (0, 1) for between survival endpoints and uniform on (-1, 0) for between survival endpoints and response. A sensitivity analysis was performed by increasing the spread of the prior mass by approximately sevenfold for the priors for treatment effects and between study variances. Adverse event rates were summarized separately along with Wilson confidence intervals for each trial [109].

3.2.2 Sequencing Assessment

Sequencing of enzalutamide and AA was assessed in the context of exponential survival models, informed by Bayesian meta-analyses with between and within study variance components. First, OS and PFS time to event data was extracted from published Kaplan-Meier curves, along with numbers at risk and censoring times, if available. Plots

were digitized using WebPlotDigitizer [110], and a custom built R [45] code was used to extract the raw time to event dataset. Censored data could only be characterized up to an interval between actual event times if censoring times were not provided. In this case, censoring times were taken at the lower bound of possible censoring times to provide a conservative estimate of information content. Then, for each time to event dataset, an exponential distribution was fit. In particular, the (monthly) rate parameter along with accompanying standard error were estimated for each time to event dataset. These rates were then meta-analyzed using a Bayesian model with between and within study variance components. Uninformative priors were used for the mean treatment effects by placing the mean at 0 and the within-study variance at 100 on the log hazard scale. A weakly informative prior was used for the between-study variance which placed 95% of the prior probability on hazard to varying up to two-fold across studies.

For constructing the sequencing assessment, rates of interest were OS and progression in the initial AR pathway inhibitor setting and OS in the post-progression after AR pathway inhibitor setting. Notably, PFS time is the minimum of OS and progression time and, in the context of the exponential model, the rate corresponding to the progression event is the difference between the PFS and OS rates. Data sources for the sequencing assessment were both randomized and observational. Randomized controlled data from the comparative effectiveness study was utilized to inform the initial AR pathway inhibitor setting, while observational data was utilized to inform the post-progression after AR pathway inhibitor setting.

The two sequencing strategies enzalutamide then AA and AA then enzalutamide, were compared by, separately for each strategy and for each of 10,000 draws from the

posterior distribution of rates, generating 5,000 initial AR pathway inhibitor OS and progression times and 5,000 post-progression after AR pathway inhibitor OS times. For each of these 5,000 sets of times, if the initial (strategy specific) AR pathway inhibitor progression time was before the OS time, then the strategy OS time was the sum of the initial AR pathway inhibitor progression time and the post-progression after AR pathway inhibitor OS time. On the other hand, if the initial (strategy specific) AR pathway inhibitor progression time was after the OS time, then the strategy OS time was simply the initial AR pathway inhibitor OS time. Based on each of these size 5,000 datasets (one dataset for each of 10,000 posterior draws) of OS times from initiation of first AR pathway inhibitor, several metrics of comparison were computed, HRs, median survival times, and probabilities of one- and two-year survival. Each of these metrics of comparison was summarized across the posterior as posterior median along with CrI.

3.3 Results

3.3.1 Comparative Effectiveness

Characteristics and efficacy summaries of studies included in the comparative effectiveness study are summarized in Table 4.

3.3.2 Overall Survival

There was weak evidence that enzalutamide outperforms AA with prednisone in terms of OS in the predocetaxel setting with posterior probability enzalutamide better than AA with prednisone on average of 0.68 and posterior probability enzalutamide better than AA with prednisone in individual study of 0.64 (HR 0.91, 95% Credible Interval (CrI)

0.62–1.35, 95% Predictive Interval (PrI) 0.55–1.53). Similarly, there was weak evidence that enzalutamide outperforms AA with prednisone in terms of overall survival in the postdocetaxel setting with posterior probability enzalutamide better than AA with prednisone on average of 0.70 and posterior probability enzalutamide better than AA with prednisone in individual study of 0.66 (HR 0.90, 95% CrI 0.61–1.33, 95% PrI 0.54–1.50). Comparative effectiveness summarized in Table 5 and Figure 12.

3.3.3 Secondary Endpoints

There was strong evidence that enzalutamide outperforms AA with prednisone in terms of secondary endpoints radiographic PFS, time until PSA progression, and PSA response rate in both the pre- and post-docetaxel settings, with posterior probabilities enzalutamide better than AA with prednisone both on average and in individual studies exceeding 0.97. Comparative effectiveness summarized in Table 5 and Figure 12.

Rates of adverse events grade ≥3 for enzalutamide versus placebo, pre-docetaxel were 46% versus 37% ($P = 0.001$), for abiraterone/prednisone versus placebo/ prednisone, pre-docetaxel were 48% versus 42% ($P = 0.057$), for enzalutamide versus placebo, postdocetaxel were 45% versus 53% ($P = 0.012$), and for abiraterone/prednisone versus placebo/prednisone, post-docetaxel were 23% versus 19% ($P = 0.146$). All reported adverse events are summarized in Table 18 in Appendix B.

3.3.4 Sensitivity Analysis

The results of the sensitivity analysis were broadly similar to the main analysis in both the pre- and post-docetaxel setting, with strong evidence of benefit for enzalutamide

relative to AA with prednisone in terms of secondary endpoints radiographic PFS, time until PSA progression, and PSA response rate. However, evidence of benefit for enzalutamide relative to AA with prednisone in terms of OS was very weak. Sensitivity analysis results summarized in in Appendix B.

3.3.5 Sequencing Assessment

Characteristics and monthly event rates of studies included in the sequencing assessment are summarized in Table 20 in Appendix B. Unfortunately, there was not sufficient data to perform a sequencing assessment in a *pre-docetaxel* setting.

Our analysis provides evidence that in the post-docetaxel setting the AA then enzalutamide strategy may be associated with longer OS time than the enzalutamide then AA strategy. In particular, respective median survival times for the AA then enzalutamide strategy and the enzalutamide then AA strategy were estimated at 21.3 months (95% CrI 16.6-28.9) and 14.7 months (95% CI 11-21.2), with posterior HR 0.66 (95% CrI 0.43-1.17) and posterior probability AA then enzalutamide better than enzalutamide then AA of 0.94. Results of sequencing analysis are summarized in Table 6.

Table 4: Summary of studies included in comparative effectiveness study

Trial	Patient population (N)	Median follow up (month)	Treatment arms (n)	Primary endpoint		Secondary endpoints					
				Overall survival		Radiographic progression free survival		Time to PSA progression		PSA response rate[c] (RR)	
				median (month)	HR (95% CI)	median (month)	HR (95% CI)	median (month)	HR (95% CI)	RR	OR (95% CI)
PREVAIL	Progressive CRPC pre-docetaxel (N=1717)	22	Enzalutamide (n=872)	not reached	0.77 (0.67 - 0.88)	not reached	0.32 (0.28 - 0.36)	11.2	0.17 (0.15 - 0.20)	59%[d]	27.23 (16.47-45.03)
			Placebo (n=845)	31.0		3.9		2.8		5%[d]	
COU-AA-302	Metastatic CRPC pre-docetaxel (N=1088)	27.1	Abiraterone acetate[b] (n=546)	35.3	0.81 (0.70 - 0.93)	16.5	0.52 (0.45 - 0.61)	11.1	0.50 (0.43 - 0.58)	68%	5.38 (4.15-6.97)
			Placebo[b] (n=542)	30.1		8.2		5.6		29%	
AFFIRM	Progressive CRPC post-docetaxel (N=1199)	14.4	Enzalutamide[a] (n=800)	18.4	0.63 (0.53 - 0.75)	8.3	0.40 (0.35 - 0.47)	8.3	0.25 (0.20 - 0.30)	54.0%[e]	76.41 (31.22-187.04)
			Placebo[a] (n=399)	13.6		2.9		3.0		1.5%[e]	
COU-AA-301	Metastatic CRPC post-docetaxel (N=1195)	20.2	Abiraterone acetate[b] (n=797)	15.8	0.74 (0.64 - 0.86)	5.6	0.66 (0.58 - 0.76)	8.5	0.63 (0.52 - 0.78)	29.5%	7.15 (4.53-11.28)
			Placebo[b] (n=398)	11.2		3.6		6.6		5.5%	

[a]Concomitant administration with prednisone was allowed but not required;
[b]concomitant administration with prednisone;
[c]response rate defined as PSA decline $\geq$ 50%;
[d]denominator (enzalutamide) was 396, denominator (placebo) was 381;
[e]denominator (enzalutamide) was 731, denominator (placebo) was 330.

Table 5: Meta-estimates for enzalutamide vs. abiraterone in the pre- and post-docetaxel settings

		Overall Survival	Radiographic Progression-Free Survival	Time to PSA Progression	PSA Response Rate
Enzalutamide vs. Abiraterone, Pre-Docetaxel	Posterior Median Hazard Ratio	0.91 (95% CrI 0.62-1.35, 95% PrI 0.55-1.53)	0.61 (95% CrI 0.41-0.91, 95% PrI 0.37-1.03)	0.37 (95% CrI 0.24-0.55, 95% PrI 0.22-0.63)	5.91 (95% CrI 3.08-11.3, 95% PrI 2.83-12.28)[b]
	Posterior Probability Hazard Ratio < 1[a]	0.68 (0.64)	0.99 (0.97)	1.00 (1.00)	1.00 (1.00)[b]
Enzalutamide vs. Abiraterone, Post-Docetaxel	Posterior Median Hazard Ratio	0.90 (95% CrI 0.61-1.33, 95% PrI 0.54-1.50)	0.61 (95% CrI 0.41-0.91, 95% PrI 0.36-1.03)	0.37 (95% CrI 0.25-0.56, 95% PrI 0.22-0.64)	6.21 (95% CrI 3.02-12.66, 95% PrI 2.79-13.65)[b]
	Posterior Probability Hazard Ratio < 1[a]	0.70 (0.66)	0.99 (0.97)	1.00 (1.00)	1.00 (1.00)[b]

[a]Posterior probability (predictive probability), [b]Odds ratio for response.

Table 6: Sequencing assessment for AA then enzalutamide strategy and the enzalutamide then AA strategy in a post-docetaxel setting

	AA then Enzalutamide	Enzalutamide then AA
Posterior HR	0.66 (95% CI 0.43-1.17)	-
Posterior Prob. HR < 1	0.94	-
Median Survival Time (Months)	21.3 (95% CI 16.6-28.9)	14.7 (95% CI 11-21.2)
Posterior Prob. One Year Survival	0.71 (95% CI 0.62-0.79)	0.58 (95% CI 0.47-0.7)
Posterior Prob. Two Year Survival	0.45 (95% CI 0.34-0.57)	0.29 (95% CI 0.1-0.45)

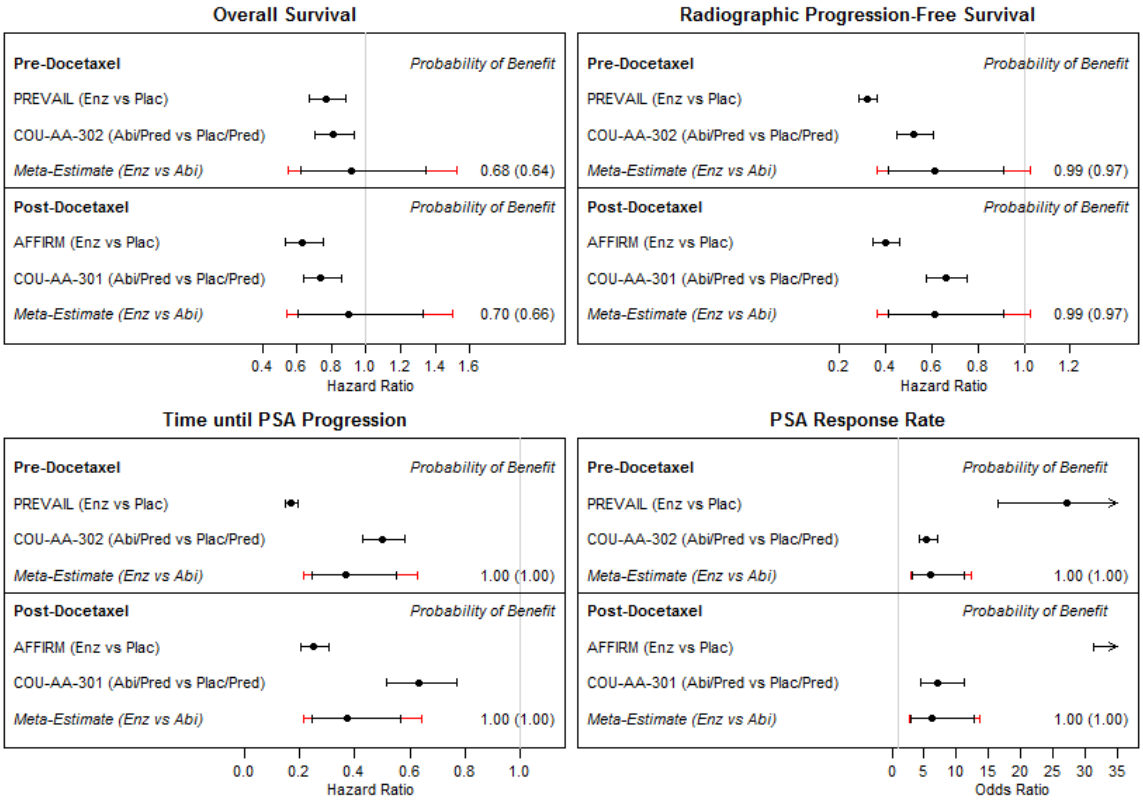

Figure 12. Individual study estimates and comparative meta-estimates for efficacy outcomes for enzalutamide vs. abiraterone in the pre- and post-docetaxel settings

3.4 Discussion

Recent availability of multiple effective agents tested in randomized trials has added complexity to the decision-making algorithm in advanced prostate cancer. The lack of randomized head-to-head comparison data between AA and enzalutamide makes it difficult to choose the optimal first-line treatment either pre- or post-chemotherapy in patients with advanced or metastatic CRPC.

We found that there is only weak evidence that enzalutamide is better than AA with prednisone in terms of OS in both the pre- and post-docetaxel setting. However, we found

strong evidence that enzalutamide outperforms AA with prednisone in terms of secondary endpoints radiographic PFS, time until PSA progression, and PSA response rate in both the pre- and post-docetaxel settings. Grade 3 or worse adverse event rates were similar between AA and enzalutamide in a pre-docetaxel setting, while there was some evidence that AA may have a lower grade 3 or worse adverse event rate than enzalutamide in a post-docetaxel setting (see Table 18). Therefore, clinicians could consider enzalutamide over AA and steroid when looking for a robust PSA response, improvement in PFS and time to PSA progression. The choice of either drug should also be tailored based on patient preferences, requirement of concomitant administration of steroids, co-morbidities and drug accessibility.

A sequencing assessment of available published trials provided some evidence that in a post-docetaxel setting, AA then enzalutamide upon progression may be associated with longer OS time than enzalutamide then AA upon progression. This finding in the sequencing assessment seems contradictory to the comparative effective analysis. There are a few potential explanations. First, there is the possibility that treatment with enzalutamide may adversely impact subsequent effectiveness of AA more than AA treatment adversely impacts subsequent treatment with enzalutamide. Another possibility is that, since evidence suggests that enzalutamide may extend time to progression without extending survival time, patients given initial enzalutamide are more likely to die before switching to AA. On the other hand, evidence suggests that patients initiated on AA are more likely to progress, and be switched to enzalutamide, before dying. Still another possibility is that the utilization of non-randomized data has led to an incorrect conclusion. Further evidence on sequencing is needed.

Our analysis becomes even more relevant with recent publication of three randomized clinical trials that suggest that androgen deprivation therapy (ADT) plus early docetaxel-based chemotherapy improves progression-free and overall survival in men with metastatic castration-sensitive prostate cancer compared with androgen deprivation therapy alone in high risk patients [111-113]. Enzalutamide may perform better as compared to AA with prednisone in these patients upon progression.

There is also some evidence from a number of small retrospective cohort studies suggesting limited activity of these agents when used in a sequential fashion either before or after docetaxel chemotherapy in advanced CRPC [97-107]. Development of predictive biomarkers to facilitate the selection of patients for a specific therapy or sequence of therapies is the focus of ongoing efforts. Recently, the AR-V7 splice variant, a truncated isoform of the AR that lacks the binding domain of both enzalutamide and AA, was shown to be associated with resistance to both agents as evidenced by inferior PSA50 response rates, PFS, and OS [108]. Prospective combination and sequence studies using both these active agents to target the androgenbased pathway in advanced CRPC are ongoing [78, 114, 115]. Data from a small prospective phase I/II study in 60 men was presented during a recent ASCO meeting [116]. Preliminary results show safety and no untoward pharmacokinetic interactions of this combination.

The comparative effectiveness study and sequencing assessment presented here have a number of limitations and strengths. Both are limited by the indirect nature of comparisons between AA and enzalutamide, which rely on the quality of between study variance component estimates and a lack of systematic interaction between individual study characteristics and treatment efficacy. Further, lack of individual patient data

precludes detailed identification of sources of study-to-study heterogeneity. The sequencing assessment is further limited by the inclusion of observational data, for which a causal link between treatment and outcomes cannot be established concretely due to various forms of confounding. On the other hand, all included studies were well-executed with objectively defined endpoints. There was no evidence of publication bias. The Bayesian approach is well-adapted to choosing between treatments with no a priori preference. In particular, the posterior probability that one treatment is better than the other summarizes the chance, conditional on the data, that one treatment outperforms the other. The posterior hazard ratio estimates, along with confidence and predictive intervals, indicate how much better might be reasonably expected. In contrast, a traditional hypothesis testing perspective is biased towards the null hypothesis, in a sense. In the absence of evidence strong enough to refute a pure sceptic, the null hypothesis is selected. On the whole, this comparative effectiveness study represents a high-quality synthesis of best-available evidence on the comparison of first-line AA and enzalutamide.

3.5 Conclusions

We sought to compare indirectly the effectiveness of abiraterone acetate and enzalutamide in advanced CRPC. Our results show that in a pre-docetaxel setting, enzalutamide may be a better drug than AA with prednisone in terms of radiographic PFS, time until PSA progression, and PSA response rate. In a post-docetaxel setting, comparative effectiveness analysis showed that enzalutamide may outperform AA with abiraterone in terms of secondary end points. The results of our analyses may help guide clinicians in making best treatment decisions with their patients. Prospective randomized

trials are eagerly awaited to provide insight on the optimal treatment sequence and combinations in this patient population.

CHAPTER 4. COST-EFFECTIVENESS OF PEMBROLIZUMAB AS FIRST-LINE THERAPY FOR ADVANCED NON-SMALL CELL LUNG CANCER

4.1 Introduction

In the last two decades, systemic therapy has brought meaningful clinical improvements for non-small cell lung cancer (NSCLC) patients, more than doubling life expectancy of patients with metastatic disease. Precision medicine and targeted therapy have become a reality responsible for increasingly high response rates and prolonged disease control for carefully selected patients [117-121]. Notwithstanding these gains, lung cancer remains the most common cause of cancer-related death, claiming more lives than breast, prostate, and colon cancer combined [122-124]. Patients lacking actionable targets - the majority - or those who inevitably progress after personalized therapy still rely upon palliative chemotherapy, with median overall survival not exceeding 16 months [123, 125, 126].

New therapies are urgently required and immunotherapy has shown enormous potential to further improve prognosis for lung cancer patients. With elevated neo-antigen expression and active mechanisms of immune surveillance evasion, lung cancer is an ideal setting for current PD-1, PD-L1, and CTLA-4 therapies [127, 128]. Recently, several PD-1, PD-L1, and CLTA-4 drugs have reached late phase development for lung cancer, in a quest for betterment of prognosis and patient selection [21, 129-131].

Some immunotherapies have received FDA and EMA approval in record time due to strong clinical results, with superior, and for some patients durable, survival and more tolerable side effects. These results have largely reset standard management of advanced NSCLC. Nevertheless, there is a price tagged to these breakthrough treatments that cannot be overlooked [132].

Analysis of cost-effectiveness of new therapies is imperative to ensure appropriate and sustainable use of advanced targeted treatments in NSCLC. The current study investigates cost-effectiveness of pembrolizumab treatment for previously untreated patients with advanced NSCLC and PD-L1 expression in ≥50% of tumor cells.

4.2 Materials and Methods

4.2.1 Data

Overall survival (OS) and progression-free survival (PFS) time-to-event data were extracted from published Kaplan-Meier (KM) curves, numbers at risk, and censoring times from the KEYNOTE-024 study. In brief, KEYNOTE-024 compared platinum-doublets versus pembrolizumab as first-line therapy for EGFR wild-type, ALK non-translocated, chemo-naive advanced lung cancer patients whose tumors expressed PD-L1 in ≥50% of cancer cells [20, 133]. The most up to date KM curves [133] for OS and PFS were digitized using WebPlotDigitizer [110], and raw time-to-event data was recovered by inverting the KM equations with a custom-built R code, extending techniques in Guyot et. al [134]. Details on frequency and severity of side effects for both intervention arms were also abstracted. We examined the quality of data recovery.

4.2.2 Bayesian survival and progression model

The distributions of OS and PFS times were modeled using a Bayesian semi-parametric framework. Specifically, we modeled discrete-time event probabilities across the time-line up to last follow-up time using a hierarchical Dirichlet Process (DP). The tail of the survival distribution (after last follow-up time) was modeled using a Weibull distribution.

The non-parametric nature of the Dirichlet process is well-suited for modeling the survival function as survival distributions are in general unlikely to follow a parametric family [135]. Bayesian nonparametric inference offers relatively new class of methods. There has been an increased interest in nonparametric approaches to analyzing survival distributions due to their considerable degree of flexibility compared to parametric alternatives. The development of Markov chain Monte Carlo (MCMC) techniques has further facilitated the success of Bayesian nonparametric inference.

One of the seminal works in the field is a paper published in *Annals of Statistics* in 1974 by Ferguson [136], who proposed the Dirichlet process as an approach to analyzing nonparametric problems from a Bayesian viewpoint. An important result is that if a sample is obtained from a mixture of Dirichlet processes, the posterior distribution of the process is again a mixture of Dirichlet processes [137]. Later in 1976, Susarla and Van Ryzin [138] initiated the modern day nonparametric Bayesian analysis of survival data in medical studies with right censored observations. They derived the Bayes estimator of the survival function under the Dirichlet process prior. Blum and Susarla [139] showed that the posterior survival distribution is a mixture of Dirichlet processes.

In 2006, Teh et al [140] extended the ordinary Dirichlet process for accommodating and modeling heterogenous groups of data. They introduced the hierarchical Dirichlet process. The model pools directly samples of survival data arising from different heterogenous groups, and lets clusters flexibly borrow information across groups. This approach avoids assumptions about the shape of the survival distributions, allows incorporation of censored data, and accommodates study-to-study heterogeneity. Furthermore, the DP model is particularly well-suited to situations where relative efficacy measures such as hazard and odds ratios are not sufficient, such as cost-effectiveness analyses requiring patient-level disease trajectories as a basic ingredient.

The DP model was constructed by modeling each *observed* event time as a multinomial variable indicating the time interval within which the corresponding event occurred. Time up to last follow-up was discretized into one-month intervals. The parameters (probability event occurs within each time interval) underlying the multinomial distributions were modeled as a study-specific Dirichlet distribution. *Right-censored* observations were also modeled as multinomial variables, with several tail categories and corresponding parameters aggregated. For example, if an observation was right-censored at 10 months, then it is known that the event occurred at some time after 10 months. The corresponding multinomial distribution would have all the categories >10 months combined into a single category, with the probability of the combined category equal to the sum of its component probabilities.

The *study-specific* Dirichlet distribution describing probabilities of events in each time interval was in turn modeled as a deviation from an *overall* Dirichlet distribution. In particular, the study-specific time interval probabilities equaled the overall Dirichlet

probabilities on average, but with study-to-study heterogeneity controlled via a scaling parameter. Weakly informative prior distributions were set for the overall Dirichlet distribution and study-to-study heterogeneity scaling parameter. Prior distributions summarize uncertainty about model parameters *before* examining the data.

In this setup, the time-line up to last follow-up was partitioned into one-month time intervals. In particular, the follow-up time for OS was 33 months and thus we partitioned the time-line into 34 one-month grids as follows:

$$[0, \text{month } 1), [\text{month } 1, \text{month } 2), \dots, [\text{month } 32, \text{month } 33), [\text{month } 34, \infty).$$

Similarly for PFS with follow-up time of 18 months, we partitioned the time-line into 19 one-month grids.

We defined the study-specific distribution, F_1, as a DP with a scaling parameter α_0, that governed the study-to-study heterogeneity, and a common base distribution, F_0. To capture the uncertainty about this distribution, we let F_0 itself be a draw from a DP with a base distribution measure S_0 that governed the a priori distribution over the data, and a concentration parameter c that captured the amount of variability around the prior S_0. This model is described as follows:

$$F_1|\alpha_0, F_0 \sim DP(\alpha_0, F_0)$$

$$F_0|c, S_0 \sim DP(c, S_0)$$

Let P_0 and P_1 correspondingly reflect the overall base probability and the study-specific probability of an event occurring in each grid on the time-line. Using the fact that

the Dirichlet distribution serves as an approximation to DP, these discrete time event probabilities follow G −dimensional Dirichlet distributions described as follows:

$$P_0(\text{grid } 1), \dots, P_0(\text{grid } G) \sim Dirichlet\ (c[S_0(\text{grid } 1), \dots, S_0(\text{grid } G)])$$

$$P_1(\text{grid } 1), \dots, P_1(\text{grid } G) \sim Dirichlet\ (\alpha_0(P_0(grid\ 1), \dots, P_0(\text{grid } G))$$

By the multinomial-Dirichlet conjugacy, event and right-censored times then follow a multinomial distribution. The observed event times, z_i, have a support on the time-line up to 1 month after last follow-up. The censored times, t_k, still have a support on the same time-line, but they were sampled over the grids on the time line until the time of censoring with the grids after that collapsed into a single tail category:

$$z_i \sim \text{Multinomial}\left(P_1(\text{grid } 1), \dots, P_1(\text{grid } G)\right) \quad \forall\, i \in \{1, \dots, I\}\,,$$

$$t_k \sim \text{Multinomial}\ (P_1(\text{grid } 1), \dots, P_1(\text{grid } r-1), \textstyle\sum_{m=r}^{G} P_1(\text{grid } m)) \quad \forall\, k \in \{1, \dots, K\}\,,$$

where *I* and *K* are the total number of events and censored observations correspondingly, and r is the censoring time.

Events that occurred in the tail of the distribution (after last follow-up time) were then modeled through a truncated Weibull survival model. The parametric nature of the Weibull model is well-suited to making inferences about the survival curve after the last follow-up time. The Weibull model was constructed by modeling survival times as random variables from a Weibull distribution within a Bayesian framework. The model had support on the whole time line beginning after the last follow-up time.

We used Markov chain Monte Carlo with Gibbs sampling to estimate the posterior distributions of F_0 and F_1 and the shape and scale parameters of the Weibull model in R through JAGS [45-47].

Model validation, sensitivity to survival function prior, and sensitivity to heterogeneity parameters are described and reported in Table 10, Figure 18, and Table 11, respectively. Samples from the posterior distribution were generated via Markov chain Monte Carlo (MCMC) implemented in JAGS and called via the rjags package in R [45-47]. Posterior distributions summarize uncertainty about the model after examining the data. Five MCMC chains were used with the first 10,000 iterations of each discarded while the Markov chain stabilized. Posterior inference was based on 100,000 iterations from each of the chains, thinned at a lag of 50.

For constructing patient trajectories in the pembrolizumab and chemotherapy arms, quantities of interest were posterior probabilities of death and progression within particular one-month time intervals up to the last follow-up, and posterior probabilities of death and progression on the continuous truncated time-line after the last follow-up time. Notably, the probability that a progression event occurred in a particular time interval was taken as the difference between the corresponding PFS and OS probabilities. Pembrolizumab and chemotherapy were compared based on 10,000 OS and progression times *for each* of the 10,000 draws from the posterior distributions of probabilities.

4.2.3 Disease model

Cost-effectiveness of pembrolizumab relative to chemotherapy was assessed using simulated patient trajectories in the pembrolizumab and chemotherapy arms over a full

lifetime horizon (with one-month increments up to last follow-up and continuous time-line after last follow-up). Patients could transition from stable disease to one transition state, (1) progressive disease, and three absorbing states, (2) death, (3) discontinuation due to treatment-related adverse events, or (4) discontinuation upon progression. Probabilities for treatment discontinuation due to adverse events and probabilities for continuation after progression were obtained from Reck et al [20, 133]. In particular, patients in the model discontinued treatment due to adverse events with probability 13.6% in the pembrolizumab group, and 10.7% in the chemotherapy group. Upon progression, 44% of patients in the pembrolizumab group, and 54% of patients in the chemotherapy group underwent second-line treatment. From a progressive disease state, patients could transition to absorbing states, (1) death or (2) treatment discontinuation. We assumed post-progression therapy discontinuation occurred after a median of 4 cycles for the pembrolizumab arm, and a median of 5 cycles for the chemotherapy arm. The state-transition diagram in Figure 13 illustrates how patients flowed through the model.

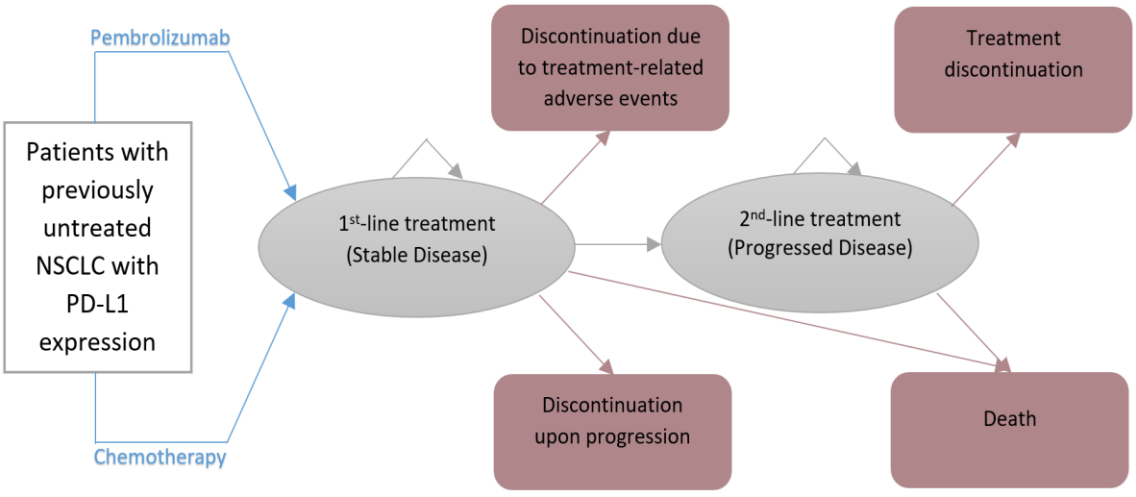

Figure 13: Patients flow

4.2.4 Dependency model

Our analysis explored several levels of dependency between each simulated patient's hypothetical disease trajectories in the pembrolizumab and chemotherapy arms, as well as between their progression and OS times. Intuitively, we might expect that a patient with longer time to progression would also have a longer survival time, and a patient with extended survival on pembrolizumab might also have a longer than typical survival time had they instead been treated with chemotherapy. The dependency model controlled the extent to which these event times were positively associated. Dependencies between each simulated patient's four associated event times (progression and death for each of pembrolizumab and chemotherapy) were modeled via a Gaussian copula.

We used a multivariate normal distribution to model the dependency between the pembrolizumab and chemotherapy arms and the dependency between progression and overall survival. We first generated four variables from a multivariate standard normal distribution with added correlation coefficients to capture the above-mentioned dependency:

$$\begin{bmatrix} z_{OS_C} \\ z_{OS_P} \\ z_{Prog_C} \\ z_{Prog_P} \end{bmatrix} \sim Normal\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & p_1 & p_2 & p_1p_2 \\ p_1 & 1 & p_1p_2 & p_2 \\ p_2 & p_1p_2 & 1 & p_1 \\ p_1p_2 & p_2 & p_1 & 1 \end{bmatrix}\right),$$

where p_1 is the correlation between the pembrolizumab and chemotherapy arms, p_2 is the correlation between overall survival and progression, $OS_C, OS_P, Prog_C$ and $Prog_P$ indicate death time for chemotherapy and pembrolizumab, and progression time for chemotherapy and pembrolizumab respectively. The death and progression times for

chemotherapy and pembrolizumab were then recovered by finding the bin (from the DP model) or the time point (from the Weibull tail) corresponding to $\Phi(z)$, where Φ is the cumulative distribution of a standard normal distribution:

$$\begin{bmatrix} \text{Death time (chemotherapy)} \\ \text{Death time (pembrolizumab)} \\ \text{Progression time (chemotherapy)} \\ \text{Progression time (pembrolizumab)} \end{bmatrix} = \begin{bmatrix} \Phi\left(z_{\mathrm{OS_C}}\right) \\ \Phi\left(z_{\mathrm{OS_P}}\right) \\ \Phi\left(z_{\mathrm{Prog_C}}\right) \\ \Phi\left(z_{\mathrm{Prog_P}}\right) \end{bmatrix}$$

Two scenarios, no and moderate dependency between hypothetical outcomes, are reported in the chapter. A high dependency scenario is described and reported in Appendix C.2.

4.2.5 Cost Data

Cost data for pembrolizumab and chemotherapy were based on UK and US costs of several aspects of treatment, care, and testing [141, 142]. The model included one-off costs for (1) PD-L1 testing, (2) enrolling under pembrolizumab therapy, (3) treatment initiation, resources upon progression specific to either (4) next line of treatment or (5) no further anti-cancer treatment, and (6) terminal care. Only those patients who entered the death state accumulated costs for end-of-life care. The model included weekly costs for resource use specific to (1) stable disease and (2) progressive disease. Treatment costs included all drugs used in first and second lines of therapy, (1) pembrolizumab (per 3-week cycle), (2) nivolumab (per 2-week cycle), and four platinum-based chemotherapy regimens per 3-week cycles: (3) carboplatin plus pemetrexed (CARB+PEM), (4) carboplatin plus paclitaxel (CARB+PAC), (5) cisplatin plus gemcitabine (CIS+GEM), and (6) pemetrexed maintenance (PEM Maint). Patients in the chemotherapy arm received one of four first-

line treatments, as reported in Reck et al [20, 133]. 30.7% received CARB+PEM followed by PEM maintenance, 37.3% CARB+PEM with no maintenance, 20.7% CIS+GEM, and 11.3% CARB+PAC. Patients in the pembrolizumab arm received pembrolizumab as first-line treatment. Patients in the chemotherapy arm who received a post-progression treatment received either (1) pembrolizumab (88%), or (2) nivolumab (12%). Assignment of post-progression treatments for the pembrolizumab arm was taken as equivalent to first-line chemotherapy arm. All costs were converted to approximate 2018 US dollars, and can be found in Table 7.

Toxicity managements costs[143-146] (Table 8) were included for several of the most common treatment-related and immune-mediated adverse effects. Hospitalization rates were obtained from Reck et al [20]. Adverse effects that were included were nausea, anemia, fatigue, diarrhea, neutropenia, vomiting, stomatitis, increased blood creatinine level, decreased platelet count, thrombocytopenia, decreased white-cell count, hypothyroidism, hyperthyroidism, and pneumonitis.

Table 7: UK and US costs of treatments, resources and tests for comparison of platinum doublet chemotherapy to pembrolizumab as first-line therapy for advanced NSCLC. All costs are in 2018 US dollars. Dosage for pembrolizumab is 200 mg every 3 weeks and for nivolumab 240 mg every 2 weeks.

Treatment/ Care/ Tests	Frequency	UK cost	US cost
Pembrolizumab (with one infusion added)	per 3-week cycle	7,558	8,760
CARB+PEM (with one infusion added)*	per 3-week cycle	2,451	6,180
CARB+PAC (with one infusion added)*	per 3-week cycle	1,322	638
CIS+GEM (with two infusions added)*	per 3-week cycle	1,294	845
PEM MAINT (with one infusion added)*	per 3-week cycle	2,226	5,983
Nivolumab	per 2-week cycle	3,896	5,926
PD-L1 single test	one-time	57	60
Total PD-L1 costs for enrolling under pembrolizumab therapy	one-time	472	-
Resource for treatment initiation	one-time	1,023	1,000
Resource use for progression-free health states	weekly	93	419
Resource upon progression to next line of treatment	one-time	1,023	1,000
Resource upon progression (no further anti-cancer treatment)	one-time	386	-
Resource use for progressed disease health state	weekly	98	-
Resource for terminal care	one-time	5,261	8,632

*Assuming body surface area of 1.80 m^2, creatinine clearance of 80 mL/min/1.73 m^2, and £150 per infusion, chemotherapy costs in UK.

Table 8: UK and US costs of adverse events for comparison of platinum doublet chemotherapy to pembrolizumab as first-line therapy for advanced NSCLC. All costs are in 2018 US dollars

	UK cost[143, 144]		US cost[145, 146]	
	Grade <3	Grade ≥3	Grade <3	Grade ≥3
Nausea	180	1,365	1,965	19,341
Anemia	3,270		4,353	20,260
Fatigue	2,902		16,185	
Diarrhea	555	1,365	3,265	16,510
Neutropenia	225		5,321	17,181
Vomiting	180	2,553	895	16,899
Constipation	-		2,591	20,949
Stomatitis	144		1,695	18,151
Increased blood creatinine level	-	729	-	-
Decreased platelet count	-	1,212	-	-
Thrombocytopenia	-	1,212	6,325	22,698
Decreased white-cell count	-			-
Dysgeusia	-		3,700	23,187
Hypothyroidism	610		2,255	20,428
Hyperthyroidism	610		2,255	20,428
Pneumonitis	2,214		9,941	21,929
Infusion reaction	-		4,782	22,860
Severe skin reaction	143		940	15,709
Colitis	1,266		6,079	20,208
Pancreatitis	-		15,943	32,918

4.2.6 Effectiveness

Patient-specific health utility values for each disease state (stable disease and disease progression) were taken according to distributions consistent with a UK-based study on health utilities for advanced NSCLC treated with immunotherapy or chemotherapy [147]. Utility values measure overall health and quality of life associated with each disease state.

Utility scores range from 1 (perfect utility) to 0 (death). Utility distributions are shown in Table 9.

The British National Institute of Clinical Excellence (NICE) provides a framework for cost-effectiveness analysis. NICE recommends that end-of-life interventions that meet the end-of-life (EoL) threshold should be given perfect utility [141]. The NICE EoL threshold is constituted of (1) small patient population, (2) prognosis <24 months, and (3) life-extension >3 months [148]. In the EoL adjusted analysis, a perfect utility was assigned to treatment with pembrolizumab if it extended life >3 months as compared to standard chemotherapy for the particular simulated patient.

Quality-adjusted life-years (QALYs) were used to measure effectiveness. QALYs are a measure combining information on both quality of life and life expectancy. QALYs are calculated as a product of time spent in each state and the corresponding health utility, and they reflect a patient's accumulated utility over time [149]. An analysis incorporating 3% annual cost and utility discounting is described and reported in Appendix C.3.

Table 9: Patient-specific utility distributions for advanced NSCLC treated with immunotherapy or chemotherapy

State	Adjusted utility
Stable disease	Uniform (0.563,0.743)
Disease progression	Uniform (0.383,0.563)

4.2.7 Analysis

Incremental cost-effectiveness ratios (ICERs) were calculated to evaluate cost-effectiveness of pembrolizumab compared to standard chemotherapy. ICERs measure incremental cost per QALY gained. A treatment is commonly considered cost-effective if the ICER is below the GBP 30,000 threshold (approximately USD 42,048) or the USD 100,000 threshold [149]. A sensitivity analysis basing the full cost-effectiveness analysis on a traditional Weibull model, instead of the combination DP-Weibull models, is presented in Appendix C.4.

4.3 Results

4.3.1 Data recovery

Data recovery using graph capture and inverting KM equations was excellent. OS and PFS time-to-event data were extracted from published KM curves, numbers at risk, and censoring times from the KEYNOTE-024 study [20, 133]. The raw time-to-event data was recovered by the inverted KM equations, extending the techniques in Guyot et al [134]. Figure 14 and Figure 15 below display how the reconstructed KM curves compare to the original KM curves published in KEYNOTE-024. In the case of OS, the recovered KM curves for both chemotherapy and pembrolizumab almost completely overlap with the original KM curves confirming that the data was recovered well. In the case of PFS, the recovered overlap with the original curves with a slight difference present after month 12. The data was recovered well.

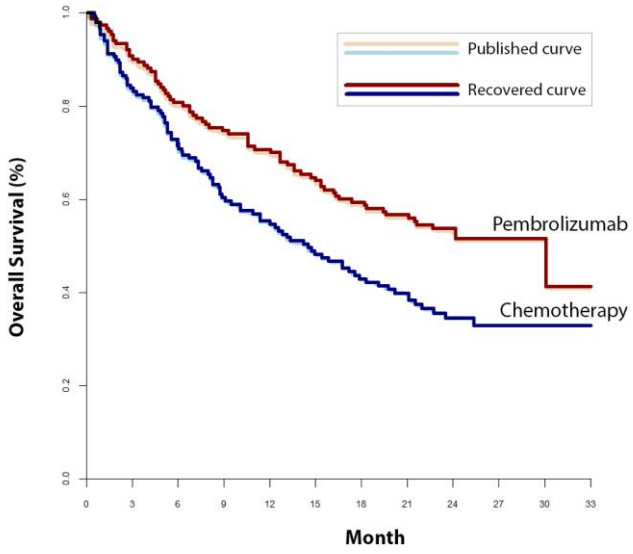

Figure 14. Reconstructed Kaplan-Meier curves for OS for the pembrolizumab and chemotherapy arms

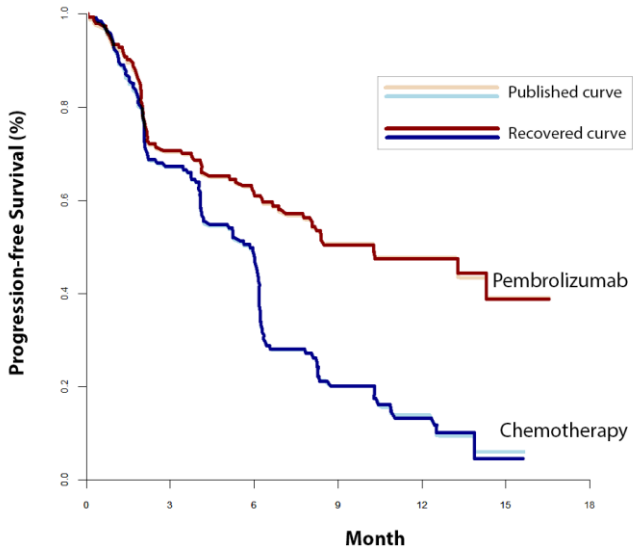

Figure 15. Reconstructed Kaplan-Meier curves for PFS for the pembrolizumab and chemotherapy arms

4.3.2 Survival model validation

The DP-Weibull models validated well, accurately reproducing several OS and PFS summary statistics (Table 10). We validated the results of our survival model by comparing posterior summary statistics to summary statistics published in KEYNOTE-024 [20, 133]. We compared the median PFS and OS survival, the OS% at 12 and 24 months and the PFS% at 6 months for both arms. Table 10 below compares the reported statistics from the study to the ones recovered by the model. The model estimates were very close to the reported statistics. Additionally, Figure 16 and Figure 17 present the model fit for OS and PFS in both arms.

Table 10: Comparison of reconstructed summary statistics to summary statistics published in KEYNOTE-024 [20, 133] – median OS and PFS, PFS% at 6 months, and OS% at 12 and 24 months for the pembrolizumab and chemotherapy arms

	Reck et al.	Model
Median PFS survival for pembrolizumab (months)	10.3	10
Median PFS survival for chemotherapy (months)	6	6
Median OS survival for pembrolizumab (months)	30	30
Median OS survival for chemotherapy (months)	14.2	15
OS % at 12 months (pembrolizumab)	70.3%	71.1%
OS% at 12 months (chemotherapy)	54.8%	54.9%
OS % at 24 months (pembrolizumab)	51.5%	51.8%
OS% at 24 months (chemotherapy)	34.5%	36.6%
PFS % at 6 months (pembrolizumab)	62.1%	62.9%
PFS% at 6 months (chemotherapy)	50.3%	49.9%

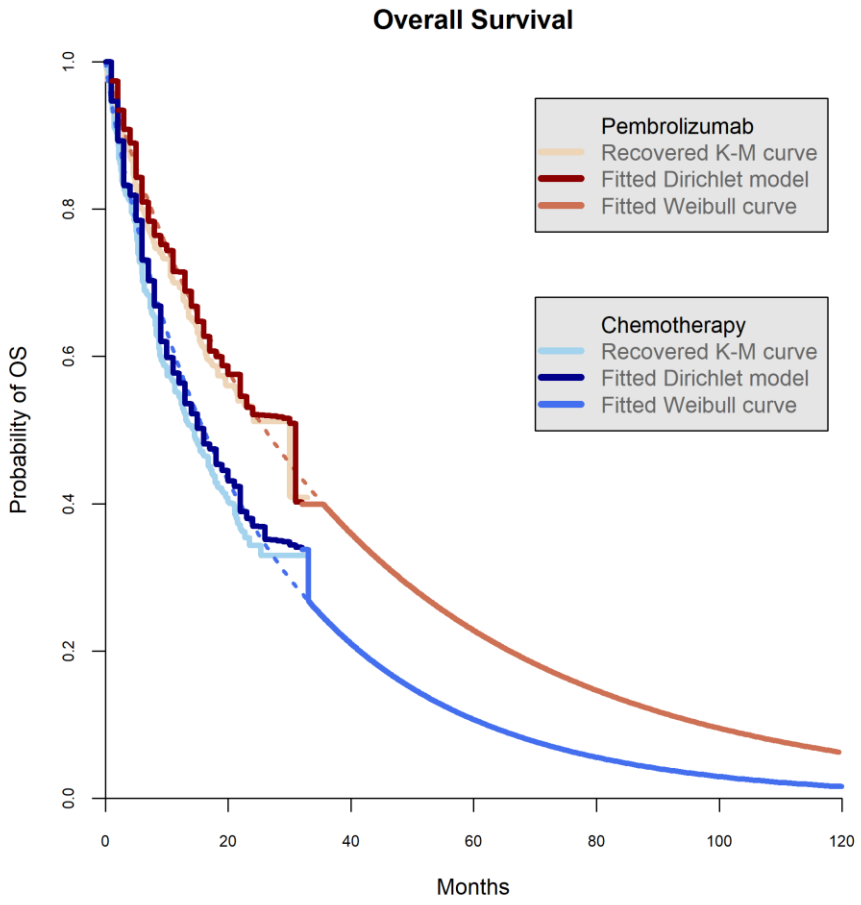

Figure 16. Reconstructed and fitted survival curves for OS for the pembrolizumab and chemotherapy arms

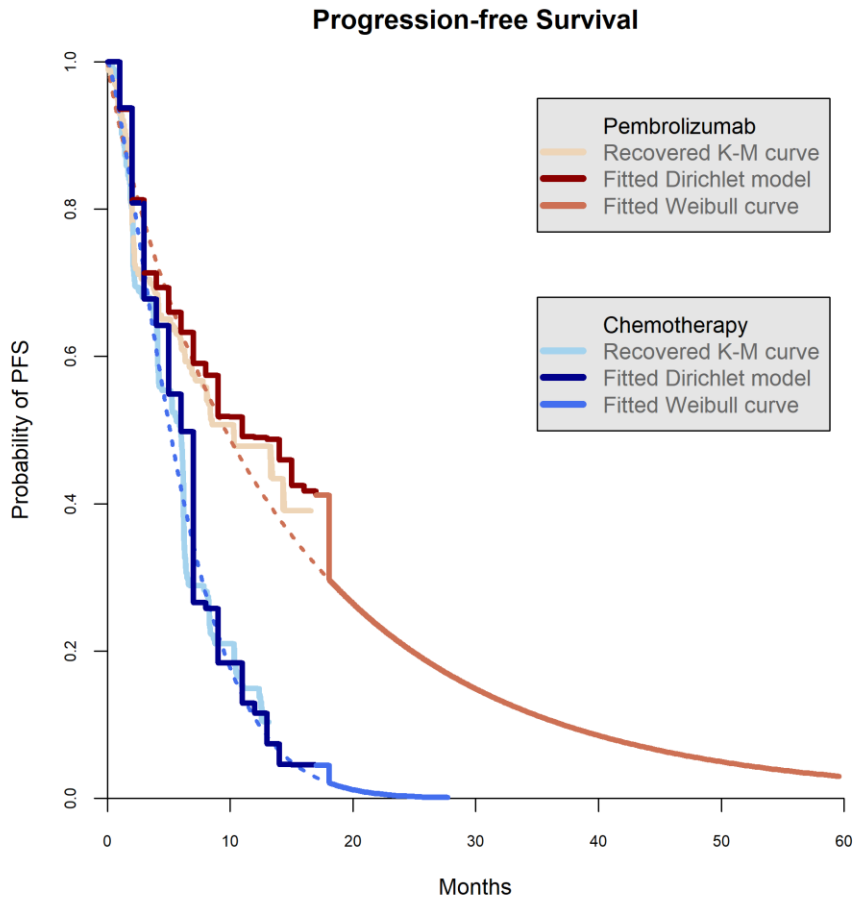

Figure 17. Reconstructed and fitted survival curves for PFS for the pembrolizumab and chemotherapy arms

4.3.3 Survival prior distribution sensitivity analysis

We explored the effect of the prior on the results by allowing the prior survival functions for both pembrolizumab and chemotherapy to vary anywhere between 0 and 1 almost everywhere across the positive axis. Figure 18 shows the spread of the prior. In the case of OS in the pembrolizumab arm, the prior survival function was allowed to vary anywhere from 0.63 to 1 at month 1 with a mean at 0.97, from 0.27 to 1 at month 2, and

anywhere from 0 to 1 for the following months. Similarly, in the case of PFS in the pembrolizumab arm, the prior was allowed to vary from 0.52 to 1 at month 1, from 0.06 to 1 at month 2, and from 0 to 1 the following months. Finally, for the chemotherapy arm, the prior for OS was spread between 0.42 and 1 at month 1 followed by a spread of 0 to 1 for the months afterwards, and for PFS it was allowed to vary from 0.18 to 1 at month 1 and then from 0 to 1 for the following months. The median OS and PFS times were 30 months (95% CrI 20-45) and 10 months (95% CrI 7-18) in the pembrolizumab arm, and 15 months (95% CrI 11-21) and 6 months (95% CrI 5-7) in the chemotherapy arm. The analysis showed that variations in priors within a sensible range had a limited qualitative effect on the outcome.

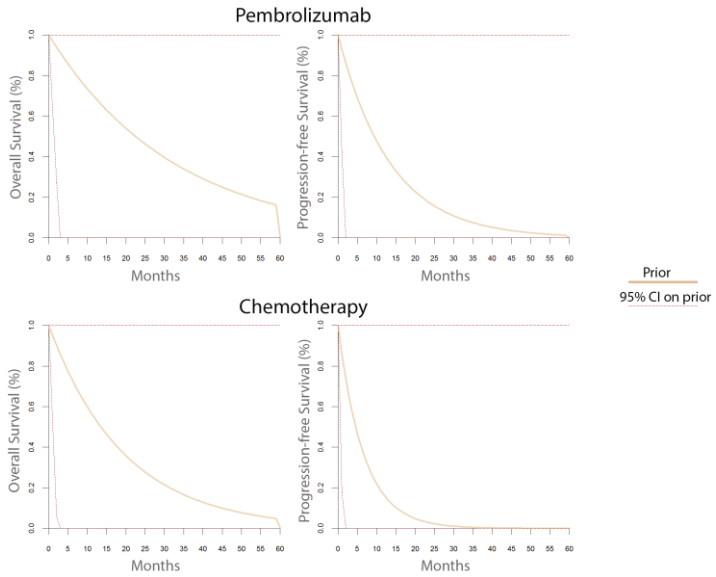

Figure 18. Prior survival function and corresponding 95% CI on the spread of the prior for OS (top left) and PFS (top right) in the pembrolizumab arm, and for OS (bottom left) and PFS (bottom right) in the chemotherapy arm.

4.3.4 Heterogeneity sensitivity analysis

We explored the effect of variable levels of study-to-study heterogeneity on the results. As previously discussed, the model is described as follows:

$$F_1|\alpha_0, F_0 \sim DP(\alpha_0, F_0)$$

$$F_0|c, S_0 \sim DP(c, S_0)$$

where the scaling parameter α_0 governs the study-to-study heterogeneity. The heterogeneity parameter was allowed to vary anywhere from 1 (introducing high level of heterogeneity) to 1,000 (almost no heterogeneity). We used the I^2 statistic to measure the magnitude of the between-study heterogeneity, where I^2 ranged from 0 (low) to 1 (high). The I^2 metric was based on variability of the median overall and progression-free survival time in the pembrolizumab and chemotherapy arms. When the initial heterogeneity was set to 1, the resulting I^2 was modestly high. For $\alpha_0 = 10$, I^2 was modest to high, and for $\alpha_0 \geq 100$, I^2 was low to modest. The results are shown in Table 11 below. We evaluated the cost-effectiveness of pembrolizumab in the case of high level of between-study heterogeneity, specifically when $\alpha_0 = 1$. While there was a significant variability present within the overall survival distribution, the study-specific survival distribution exhibited little to almost no change. This behavior shows a limitation of our analysis that stems from the absence of multiple studies. Since the model was based on a single study, the analysis showed that introducing study-to-study heterogeneity had little qualitative effect on the results (see Table 21 in Appendix C.1).

Table 11: I^2 values for different levels of study-to-study heterogeneity across OS and PFS for the pembrolizumab and chemotherapy arms. I^2 metric based on the variability of the median OS and PFS for each arm

	I^2			
	Pembrolizumab		Chemotherapy	
α_0	OS	PFS	OS	PFS
1	0.57	0.14	0.72	0.76
10	0.36	0.1	0.53	0.61
100	0.05	0.22	0.09	0.32
1000	0.03	0.02	0.06	0.01

4.3.5 Cost-effectiveness

For the no dependency among outcomes scenario, we modeled the two arms and their corresponding OS and PFS times as fully independent of each other. Patients who received chemotherapy gained a posterior mean of 1.11 QALYs (95% CrI 0.99-1.18). Patients who received pembrolizumab gained a posterior mean 1.93 QALYs (95% CrI 1.7-2.01) or EoL adjusted 3.06 QALYs (95% CrI 2.63-3.23). Posterior mean UK and US costs for the duration of therapy in the pembrolizumab arm were $99,000 (UK) and $132,000 (US), compared to $34,000 (UK) and $73,000 (US), respectively, for chemotherapy. These translated into posterior mean ICERs of $81,000 per QALY in the UK setting and $74,000 per QALY in the US setting. With EoL adjustment, the respective UK and US posterior mean ICERs were $34,000 and $31,000 per QALY (Table 12 and Figure 19).

The probability of pembrolizumab being cost-effective was <1% with respect to UK (USD 42,048) threshold without EoL adjustment, and 97.1% with EoL adjustment. The

probability that pembrolizumab was cost-effective with respect to US (USD 100,000) threshold was 97.2% without EoL adjustment and >99% with EoL adjustment (Figure 19).

In the second scenario, we incorporated a moderate dependency between each simulated patient's outcomes in the pembrolizumab and chemotherapy arms and between their associated OS and progression times by introducing a (latent) correlation of 0.5 between the pembrolizumab and chemotherapy arms and a (latent) correlation of 0.8 between progression and OS times via a Gaussian copula. In the chemotherapy arm, posterior mean QALYs gained decreased to 1.06, while mean treatment cost increased to $38,000 in the UK setting, and to $81,000 in the US setting. Mean cost in the pembrolizumab arm increased to $121,000 in the UK setting, and to $160,000 in the US setting. In the absence of EoL adjustment, mean QALYs gained by patients on pembrolizumab decreased to 1.80, leading to ICER per QALY of $115,000 for the UK and $110,000 for the US. With EoL adjustment, ICERs per QALY for the UK and US settings were $52,000 and $49,000, respectively (Table 12 and Figure 19). The probability of pembrolizumab being cost-effective in the UK was <1% with or without EoL adjustment. In the US setting, the probability that pembrolizumab was cost-effective was 25.32%, and 99.8% with EoL adjustment (Table 12 and Figure 19).

We explored a third, high dependence, scenario which incorporated a strong dependency between the arms and the progression and survival times. Results were qualitatively similar, with higher levels of dependency leading to higher ICERs, and can be found in Appendix C.2.

Figure 20 displays cost-effectiveness acceptability curves for pembrolizumab for the no and moderate dependency scenarios in a UK and US setting. In the UK setting, there is a 50% probability that pembrolizumab is cost-effective at a willingness-to-pay threshold of USD ≥51,000 per QALY in the scenario with EoL adjustment, and USD ≥111,000 per QALY without EoL adjustment. In the US setting, there is a 50% probability that pembrolizumab is cost-effective at a willingness-to-pay threshold of USD ≥48,500 per QALY in the scenario with EoL adjustment, and USD ≥106,500 per QALY without EoL adjustment.

Additionally, a discounting factor was considered, but results were qualitatively similar with slightly higher ICER values (Appendix C.3). Finally, a relatively traditional survival model based on the Weibull distribution was considered, and the results were qualitatively similar to the DP-Weibull model in each of the no, moderate, and high dependency scenarios (Appendix C.4).

Table 12: Posterior mean (95% CrI) costs (2018 USD), QALYs and ICERs for comparison of platinum doublet chemotherapy versus pembrolizumab as first-line therapy for advanced NSCLC, under no dependency and moderate dependency between patients' hypothetical outcomes

	Chemotherapy		Pembrolizumab				
				Without End-of-Life adjustment		With End-of-Life adjustment	
	Cost ('000) USD (95% CrI)	QALY (95% CrI)	Cost ('000) USD (95% CrI)	QALY (95% CrI)	ICER ('000) USD per QALY (95% CrI)	QALY (95% CrI)	ICER ('000) USD per QALY (95% CrI)
No Dependency							
UK	34 (32-36)	1.11 (0.99-1.18)	99 (90-108)	1.93 (1.7-2.01)	81 (67-111)	3.06 (2.63-3.23)	34 (29-43)
US	73 (68-77)		132 (120-144)		74 (58-102)		31 (25-40)
Moderate Dependency							
UK	38 (35-40)	1.06 (0.94-1.13)	121 (112-128)	1.8 (1.56-1.89)	115 (93-166)	2.69 (2.26-2.86)	52 (43-69)
US	81 (76-85)		160 (150-170)		110 (87-159)		49 (40-67)

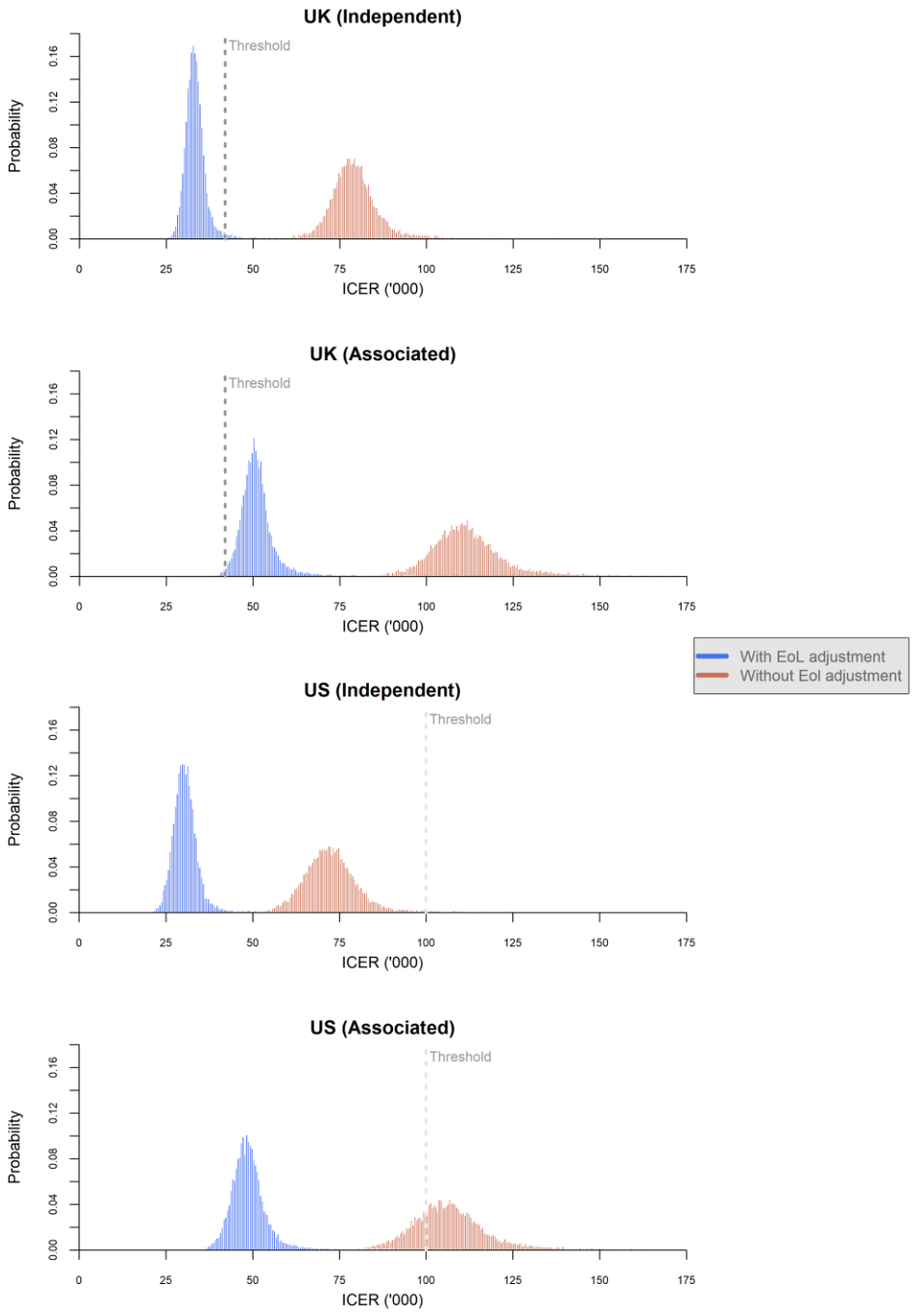

Figure 19. Posterior distribution ICER

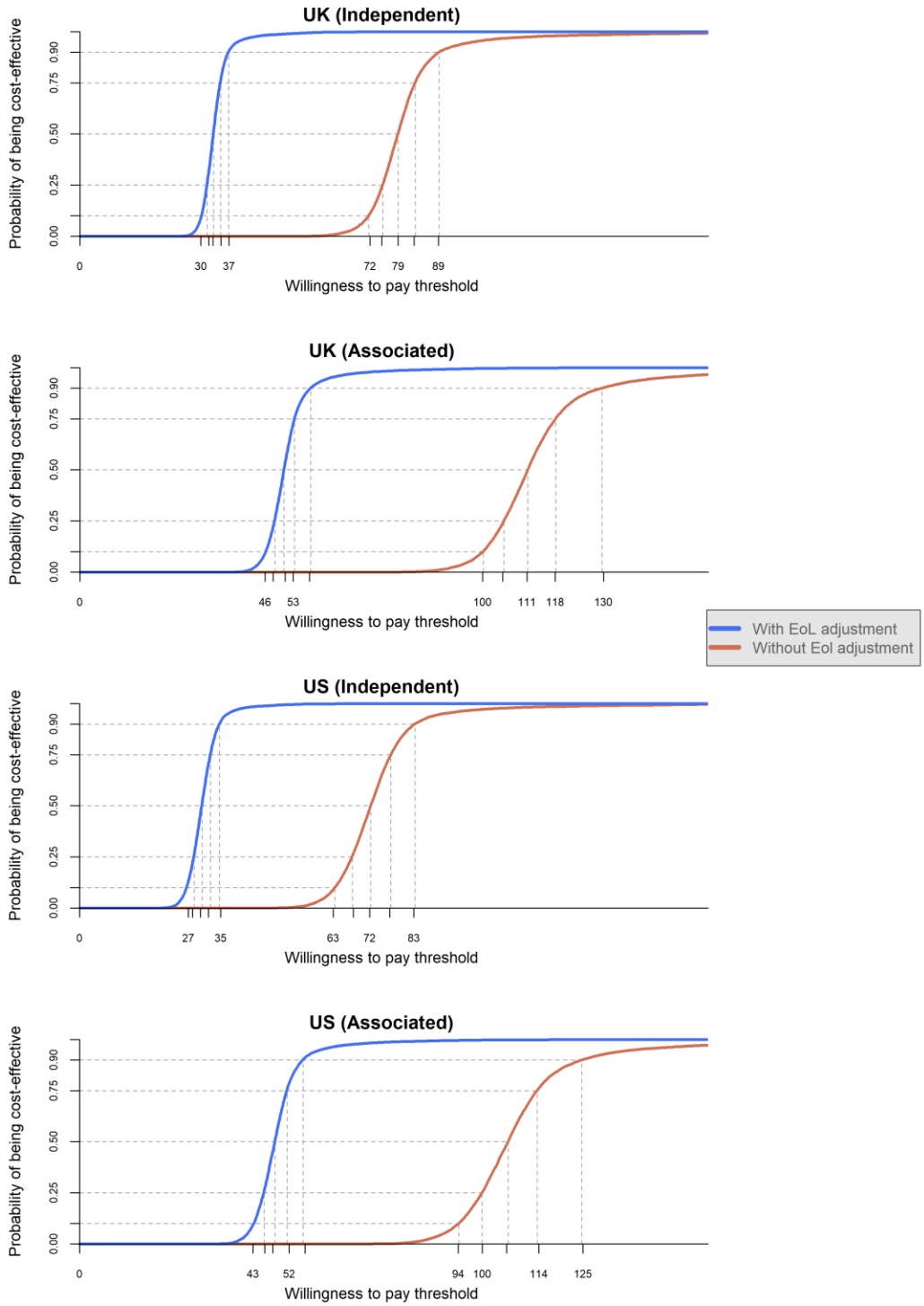

Figure 20. Acceptability ICER

4.4 Discussion

The success of immune checkpoint blockade has no parallel in recent medical oncology, however its costs are a matter of concern. Our current analysis indicates that pembrolizumab is likely to be cost-effective within the US but not in the UK. This difference across the Atlantic seems to stem from different willingness to pay thresholds (US USD 100,000, UK USD 42048), as the US and UK ICER values were close to each other in nearly all sensitivity and dependency analyses. Pembrolizumab was cost-effective in nearly all sub-analysis in US setting, whereas in the UK setting, evidence suggested it may be cost-effective only in the no dependency model, a fairly stringent assumption.

We used the end of life adjustment [149]– a tool proposed and applied by NICE to assess therapies in challenging palliative settings – and the results became even more favored towards incorporation of pembrolizumab in the US. EoL adjustment also decreased the UK ICERs, but the values fell short of the current British willingness to pay threshold. It is noteworthy that ICER values were numerically similar between US and UK in all analyses, but the difference in thresholds (USD 100,000 in the US and USD 42000 in the UK) precluded a favorable opinion in UK setting.

Moving beyond the US scenario, the enticing clinical results of pembrolizumab must be put into perspective against the backdrop of limited reimbursement supply, present even in the high-income world, as in the UK scenario. The conundrum between clinical efficacy and costs can reach dramatic proportions when a groundbreaking therapy emerges but its costs are prohibitive, which seems to be the case with first-line pembrolizumab for lung cancer in the middle and lower income countries. It is crucial to inform the general

population, patients, physicians, and paying sources about those limitations in order to best allocate scarce funds.

A pragmatic approach to bring new and expensive drugs to clinics would be lowering their prices and some examples have shown this approach is feasible. For instance, nivolumab and pembrolizumab have been incorporated as second or third line therapy for lung cancer patients in the UK after NICE successfully negotiated lower costs and proposed more stringent patient selection [150]. Those treatments had original ICERs far above the British threshold, but became available to NHS patients after aggressive deals were struck with pharmaceutical companies. Even though these transactions were productive, this path may not be generalizable worldwide. Not all patients will be represented by a stakeholder with the bargaining power of the NHS, nor will there be alternative treatment to which patients can switch in the case of unproductive negotiations.

Other proposals must also be explored. One alternative recently put forth is cost-sharing between pharmaceutical companies and the reimbursement body, as in the Italian public health system by a manufacturer of another PD-1 antibody [151]. In that specific deal, initial costs would be defrayed by the reimbursement body and later drug-only costs by the pharmaceutical company. It is paramount to notice that the main cost still falls on the reimbursement body, as fewer patients will be at the tail of therapy and all non-drug costs are paid by the health care provider.

Another approach for accepting more expensive drugs would be raising the willingness to pay bar in the specific scenarios of “end of life treatments”. This is a real alternative, already adopted by NHS for treating metastatic melanoma with ipilimumab, a

practice changing immunotherapy [152]. The current use of “end of life” adjustments supports the application of this tool in our present work. Those approaches are promising but may not suffice. A mixed option of discount, cost-sharing, end of life adjustment and multicriteria decision analysis [153] is more likely to be provide better access to high value therapies.

A unique and very relevant feature of immunotherapy is the long-term survivorship brought by this type of treatment, that may not be captured by standard statistics used in cost-effectiveness analysis or studies with immature overall survival data [154]. The real - even though small - group of long term survivors may positively impact on survival outcomes, rendering the investment in the therapy still more appealing as the update of Keynote 024 has shown [133].

Even though evidence suggests pembrolizumab may be considered cost-effective in the US, it is expected that immunotherapy costs will be a constant as newer therapies are set to be more expensive [155]. An example for the present day in the realms of lung cancer is the FDA approval of pembrolizumab combined with platinum doublet for first-line chemotherapy based on Keynote-021 [155]. Again, the clinical results are enticing, with subgroup analysis showing lasting tumor response in 80% of PD-L1 positive patients. However, it does not appear that immunotherapy-chemotherapy combination will become a cost-sustainable option even for this selected group at current costs of PD-1 therapy even in high income countries.

As in most cost-effectiveness analyses, our results and conclusions were based upon adaptations from clinical trials to the real world and further limited by lack of

comprehensive cost and survival data. In our model, patients received further treatment according to a distribution abstracted from the chemotherapy arm. However, duration and outcomes on second-line treatment are not described in detail and there was a high rate of treatment cross over, adding uncertainty to the results [156]. Nevertheless, we used the most updated Keynote 024 results with mature overall survival [133].

In spite of its limitations, our work can improve understanding of costs. Detailed costs of immune-related side-effects, both mild and severe, were retrieved and applied. Nivolumab, another expensive immunotherapy available as second or third line in the UK and US, entered the model as an option for further therapy. We ran analyses in two different scenarios: the mostly public-funded UK, and the mixed reimbursement US.

A recent analysis by Huang et al. assessed cost effectiveness of pembrolizumab compared with platinum-based chemotherapy as first-line treatment in patients with metastatic NSCLC with PD-L1 expression ≥50% at $97,621 per QALY gained for the US [157]. That analysis was based on the preliminary data from Keynote 024 study, assumed a parametric distribution for the overall and progression-free survival and relied on lengthy extrapolation to model a 20-year time horizon. The study estimated the mean time in the PFS state for pembrolizumab and chemotherapy at 2.16 and 0.55 years respectively [157]. Our analysis on the other hand, estimated the mean time spent in the PFS state for pembrolizumab and chemotherapy at less than 10 and 6 months correspondingly, which is more consistent with the respective median PFS times of 10.3 and 6 months, as reported at Reck et al [20, 133]. Moreover, we included other PD-1 therapies and platinum-doublets as possible further therapies, which very likely increased the overall costs of therapies.

Cost-effectiveness studies aim to inform limitations of a given therapy and to allow a better decision-making process. Funders, pharmaceutical companies and, most importantly, patients and families must be actively involved in these decisions bearing in mind benefits and costs of new therapies in the joint attempt of providing sustainable cancer care.

CHAPTER 5. BINOMIAL n −PROBLEM

5.1 Introduction

Because of its applicability, the problem of estimating the success probability p in a binomial $\mathrm{Bin}(n, p)$ distribution when the number of trials n is known is one of the most fundamental statistical problems finding place in introductory statistical textbooks. However, in some real-life situations, n may not be known, and may be the parameter of interest. The binomial n problem is a much less studied and a much more difficult problem.

The most difficult cases involve simultaneous estimation of both n and p, even when only a single observation might be available. The problem of estimating the parameters n and p simultaneously has been first addressed in the literature by Whitaker [158], Fisher [159] and Haldane [160], who proposed the method of moments estimators (MME's) and presented how the maximum likelihood estimates (MLE's) may be computed. Fisher [159] didn't take the problem seriously arguing that for sufficiently large number of observations k, n will be known. While this argument is correct, if p is small, in practice k will have to be unreasonably large for n to be known with any degree of certainty.

Classical procedures and their asymptotic properties were further critically assessed by Olkin et al [24], Carroll and Lombard [161], and Casella [162], who pointed out that the MME (introduced by Haldane [160]) and the MLE (introduced by Fisher [159]) estimators were highly sensitive to slight perturbations of the count data and hence unstable. This erratic behavior arises when the sample mean and sample variance of the observations are nearly equal. To overcome this lack of robustness, Olkin et al [24]

proposed two stabilized versions of the MME and MLE estimators: jackknife-stabilized MLE:S, and ridge-stabilized MME:S. Both the MLE:S and MME:S estimators outperform the ordinary MLE and MME and are reasonably stable. Later, Carroll and Lombard [161] took a different approach to stabilizing the classical estimators. They suggested an alternative estimator $\mathrm{MB}(a, b)$ by assuming a beta prior distribution for p and maximizing an integrated likelihood. Casella [162] explored situations in which stabilized estimators are preferred over classical estimators. Further classic literature includes Blumenthal and Dahiya [163], Lindsay [164], Hall [165], Kuhne [166], and others.

Recently, DasGupta and Rubin [167] proposed two new more efficient estimators. The first one is a new moment estimator that uses the sample maximum, the sample mean and the sample variance and is easy to motivate and compute. The second one is a bias correction of the sample maximum that performs very well and outperforms the Carroll-Lombard estimate in many scenarios. The authors also derived the two estimators' asymptotic properties.

While most of the classical estimators of n are based on MLE, MVUE or MME, several authors have considered the binomial n problem from a Bayesian viewpoint. Draper and Guttman [23] proposed a Bayes point estimate assuming that the prior distribution of n is discrete uniform on a set $\{1,2, \dots, N\}$ with a prespecified upper bound N. In the case when p is known, the Bayes estimator coincides with the MLE solution of Feldman and Fox [168], and in the case when p is unknown, p is assumed to follow a beta prior distribution that is independent of n. Raftery [169] adopted a Bayes empirical Bayes approach within a hierarchical under the assumption that n follows a Poisson distribution.

Kahn [170] demonstrated that the tails of the posterior distribution of n are fully determined by the prior distribution of p, and do not depend on the data. Hamedani and Walter [171] investigated Bayes estimators based on a general prior distribution for n. Gunel and Chilko [172] proposed a Bayesian estimate under a continuous prior distribution for n. More recently, Bayoud [173] proposed Bayes and empirical Bayes point estimates for n under the assumption of a left-truncated Poisson prior distribution for n and a beta prior distribution for p. Bayesian approaches seem to alleviate difficulties inherent to the classical approaches, but they do no appeal to asymptotic theory, consequently being for practical "small" problems.

The simpler case of estimating the parameter n when p is known has been addressed by Feldman and Fox [168] who obtain estimates based on MLE, MVUE and MME and develop their asymptotic properties. Hunter and Griffiths [174], Sadooghi-Alvandi [175], Zou and Wan [176], Iliopoulos [177], Bayoud [173], and De and Zacks [178].

The binomial n problem when p unknown continues to be a fundamentally difficult problem. As mentioned earlier the problem exhibits an inherent instability. Furthermore, DasGupta and Rubin [167] established that neither n nor p are unbiasedly estimable which imposes further difficulty obtaining good estimates. The most profound difficulty across estimators arises from their tendency to severely underestimate n, especially when either n is large, or p is small. Furthermore, in the absence of replication, inference about n is not possible.

The problem that we address in this paper is estimating the parameter n in a binomial distribution when p is unknown for both cases of multiple observations and no replications.

5.2 n −estimators

Given a random sample $X_1, \dots, X_k$ of iid observations each drawn from a binomial distribution $\text{Bin}(n, p)$, where n, the total number of trials, and p, the success probability, are unknown, we consider the problem of estimating n. We assume that n and p are independent, $n \in \{1,2, \dots\}$ and $p \in (0,1)$. Define the sample mean, the sample variance and the sample maximum as $\bar{x} = \frac{\sum_{i=1}^{k} x_i}{k}$, $s^2 = \frac{\sum_{i=1}^{k}(x_i - \bar{x})^2}{k}$ and $x_{(k)}$ respectively.

The n binomial problem is a notoriously difficult one. The most common issue with estimators of n is their instability. Both the MLE and MME estimators are highly sensitive to even slight perturbations of the data when the sample mean is nearly equal to or exceeds the sample variance. In order to avoid this issue, various stabilized estimators have been proposed, but a lot of them tend to underestimate n, especially when p is small. Smaller variance occurs when p is near 0 or 1. Variance is maximized at $p = \frac{1}{2}$. Given a sample with a small variance, it's hard to distinguish whether the sample comes from a population with small p and large n, or large p and small n. Most estimators tend to go for the latter and severely underestimate n, which poses a serious practical difficulty.

DasGupta and Rubin [167] established the lack of unbiased estimates.

Another difficulty is that, as noted by Student [179] and Olkin et al [24], there are certain ranges of the parameters n and p for which it is unclear whether the binomial or Poisson distribution is a better fit. If p is small and n is large, the Poisson distribution is a good approximation to the binomial distribution.

A more practical approach to the binomial n problem is provided from a Bayesian standpoint. Analytical results might not be feasible, but numerical results are usually easy to obtain. Additionally, the Bayesian approach does not require any appeal to asymptotic distributional results. However, it exhibits an inherent challenge from a Bayesian perspective because n is discrete. That restriction leads to a limited choice of prior distributions. Furthermore, Jeffrey priors are not defined. Additionally, the posterior of n must be restricted to $n \geq x_{(k)}$.

Kahn [170] showed that for large n the tail weight of the marginal posterior distribution of n is totally determined by the choice of prior distribution on n and p rather than the data. In fact, the tail of this posterior is not asymptotically affected by the data.

Finally, the Bayesian approaches pose an additional question on how one should specify the parameters of the prior distributions.

Now, we present existing estimators of the sample size n when p is unknown.

5.2.1 Estimators related to the sample maximum

5.2.1.1 Sample maximum

A trivial estimator of n is $x_{(k)}$. As $k \to \infty$, the maximum sample is a consistent estimator of n, but a biased one as it can severely underestimate n. In practice, k will have to be unrealistically large before n can be known with any degree of certainty. One can modify the estimator by adding a constant r. Feldman and Fox [168] showed that the estimator $x_{(k)} + r$ is still a consistent estimator of n as $n, k \to \infty$ provided that $kp^n n^{r-1} \to \infty$, but it is nevertheless still unreliable.

5.2.1.2 Bias correction of the sample maximum

An attempt to improve the sample maximum estimator is to incorporate a bias correction. DasGupta and Rubin [167] introduced the following bias corrected estimator by obtaining a bound on the average bias of $x_{(k)}$ averaged over p:

$$\hat{n} = x_{(k)} + \sum_{i=0}^{\tilde{n}-2} F^{-1}_{i+1,\tilde{n}-i}\left(\frac{1}{k}\right),$$

where $F^{-1}_{r,s}$ denotes the quantile function of the Beta(r,s) distribution and $\tilde{n}$ is some suitable (preliminary) estimate of n.

5.2.2 Estimators related to MME

5.2.2.1 MME

Haldane [160] presented the method of moments estimators of the parameters p and n:

$$\hat{p} = \frac{\bar{x} - s^2}{\bar{x}} \qquad \text{and} \qquad \hat{n} = \frac{\bar{x}^2}{\bar{x} - s^2}.$$

While the MME estimators are consistent, they do not seem to be fully efficient according to Haldane [160]. The MME estimator is very unstable when s^2 is close to $\bar{x}$. Furthermore, produces negative estimates with positive probability. Despite its erratic behavior especially when n is large and p is small, the MME can still be a useful estimate. Binet [180] showed that if the MLE estimator exceeds one hundred, very little information

or efficiency is lost by using the MME estimator instead. Furthermore, the MME is simple to compute.

5.2.2.2 Stabilized MME

Olkin et al [24] proposed a stabilized version of the MME estimator. In the case when s^2 is close to $\bar{x}$ and the MME is not stable, the authors suggested perturbing the observed sample by a small constant $\epsilon > 0$, so that the sample mean of the perturbed observations becomes greater than s^2. The stabilized version of MME is defined as:

$$\hat{n} = \begin{cases} \max\left\{\frac{\bar{x}^2}{\bar{x} - s^2}, x_{(k)}\right\}, & \text{if} \quad \frac{\bar{x}}{s^2} \geq 1 + \frac{1}{\sqrt{2}} \\ \left(2 + \frac{3}{\sqrt{2}}\right)s^2, & \text{if} \quad \frac{\bar{x}}{s^2} < 1 + \frac{1}{\sqrt{2}} \quad \text{and} \quad s \geq \frac{z_k}{1+\sqrt{2}} \\ \frac{z_k^2 s}{z_k - s}, & \text{if} \quad \frac{\bar{x}}{s^2} < 1 + \frac{1}{\sqrt{2}} \quad \text{and} \quad s < \frac{z_k}{1+\sqrt{2}} \end{cases},$$

where $z_k = \Phi^{-1}\left(2^{-1/k}\right)$, and Φ denotes the standard normal cumulative distribution function.

5.2.2.3 A new moment estimate

DasGupta and Rubin [167] introduced a new moment estimate, that incorporates information not only from $\bar{x}$ and s^2, but also from $x_{(k)}$:

$$\hat{n} = \frac{x_{(k)}^{\alpha+1}(s^2)^{\alpha}}{\bar{x}^{\alpha}\left(x_{(k)} - \bar{x}\right)^{\alpha}}$$

The estimator is easy to compute. The authors state that a good choice for α is 1.

5.2.3 Estimators related to MLE

5.2.3.1 MLE

The maximum likelihood estimator, MLE, is defined as:

$$\hat{n} = \begin{cases} x_{(k)}, & \text{if} \quad d\left(x_{(k)}\right) \leq 0 \\ \infty, & \text{if} \quad s^2 \geq \bar{x} \\ \text{solution of } d(n) = 0, & \text{if} \quad d\left(x_{(k)}\right) > 0 \text{ and } s^2 < \bar{x} \end{cases}$$

where $d(n) = k \log\left(1 - \frac{\bar{x}}{n}\right) + \sum_{j=0}^{x_{(k)}-1} \sum_{x=j+1}^{x_{(k)}} \frac{N_x}{n-j}$ with N_x being the sample frequency of x. The maximum likelihood equation $d(n)$ is presented by Fisher [159] and Haldane [160].Olkin et al [24] shows that the likelihood equation leads to a unique MLE estimator if $d\left(x_{(k)}\right) > 0$. Similarly to the MME estimator, as $\frac{\bar{x}}{s^2} \to 1$, the estimator become less stable and more sensitive to small changes in the success counts. Furthermore, as $n \to \infty$, it is not necessarily true that $L(n) \to \infty$, where $L(n)$ is the likelihood function, thus causing instability issues.

5.2.3.2 Stabilized MLE

Olkin et al [24] robustified the MLE estimator through a jackknife procedure and arrived at the following stabilized MLE estimator that is either the ordinary MLE estimator or a jackknifed version of $x_{(k)}$:

$$\hat{n} = \begin{cases} \text{MLE}, & \text{if} \quad \frac{\bar{x}}{s^2} \geq 1 + \frac{1}{\sqrt{2}} \\ x_{(k)} + \frac{k-1}{k}\left[x_{(k)} - x_{(k-1)}\right], & \text{otherwise} \end{cases}$$

where $x_{(k-1)}$ is the $k-1$ order statistic.

5.2.3.3 Second stabilized MLE

Casella [162] analyzed the stability of the MLE estimator by perturbing the log likelihood function in a systematic fashion. The author argues that the instability arises from the first term in the log likelihood function given below:

$$l(n,p|X,k) = \sum_{i=1}^{k} \log \binom{n}{x_i} + k\bar{x} \log p + k(n-\bar{x}) \log(1-p).$$

The goal is to perturb the first term by replacing the $\log n!$ and $\log(n-x_i)!$ terms with the following approximation:

$$\log y! \approx (1-\alpha) y \log y + \alpha(y+1) \log(y+1),$$

for some $\alpha \in (0,1)$. For α near $\frac{1}{2}$, the perturbed likelihood is very close to the original one. For a range of α values near $\frac{1}{2}$, the author examines the maximum likelihood estimators resulting from those slightly perturbed likelihood functions. Differentiating $l(n,p|X,k)$ leads to the stabilized n −estimator $\hat{n}_\alpha$, which is the solution to

$$\log \left[\frac{n^{(1-\alpha)k}(n+1)^{\alpha k} \left[\frac{\sum_{i=1}^{k}(n-x_i)}{kn} \right]^k}{\prod_{i=1}^{k}(n-x_i)^{\alpha} \prod_{i=1}^{k}(n-x_i+1)^{1-\alpha}} \right] = 0.$$

The MLE of p based on the perturbed likelihood is $\hat{p}_\alpha = \frac{\bar{x}}{\hat{n}_\alpha}$. Casella shows that for a fixed α, the perturbed likelihood has a unique finite root $\hat{n}_\alpha$. To obtain a point estimate of n, the author suggests taking a weighted average of values of $\hat{n}_\alpha$ for α near $\frac{1}{2}$:

$$\hat{n}_{\text{avg}} = \frac{\sum_\alpha \left|\frac{1}{2} - \alpha\right| \hat{n}_\alpha}{\sum_\alpha \left|\frac{1}{2} - \alpha\right|}$$

5.2.3.4 Likelihood weighed by beta prior for $\boldsymbol{p}$

Consider the conjugate $\text{Beta}(\alpha, \beta)$ prior distribution on the parameter p. The justification behind the beta distribution is that it does not impose any severe limitations on the way the probability fluctuates, yet it is versatile. The marginal posterior distribution for n after integrating out p and recognizing the beta integral is:

$$L(n) = \int_0^1 \prod_{i=1}^{k} \binom{n}{x_i} p^{x_i}(1-p)^{n-x_i} \frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha,\beta)} dp$$

$$= \frac{B\left(\alpha + \sum_{i=1}^k x_i\,, \beta + kn - \sum_{i=1}^k x_i\right)}{B(\alpha,\beta)} \prod_{i=1}^{k} \binom{n}{x_i}.$$

Carroll and Lombard [161] obtain an estimator by maximizing the marginal posterior as a function of $n \geq x_{(k)}$ for some given α and β. According to DasGupta and Rubin [167], it is the best available estimate of n.

Skellam [181] discussed how to obtain satisfactory estimates of the parameters α and β through the method of moments through the following relation:

$$\alpha = \frac{R_1 R_2 - (n-1)R_1}{(n-1)R_1 - nR_2}$$

$$\beta = \alpha\left(\frac{n}{R_1} - 1\right)$$

where $R_j = \frac{(n-j+1)(j+\alpha-1)}{j+\alpha+\beta-1}$. More efficient parameters can be obtained by directly maximizing the log likelihood instead.

Carroll and Lombard [161] examine this problem when α and β are integers. Their estimator $\mathrm{MB}(\alpha-1, \beta-1)$ is obtained by maximizing the integrated likelihood as a function of $n \geq x_{(k)}$, where

$$L(n) = \prod_{i=1}^{k} \binom{n}{x_i} \frac{1}{(\alpha+\beta+kn-1)\binom{\alpha+\beta+kn-2}{\alpha+\sum_{i=1}^{k} x_i - 1}}.$$

Notice that the likelihood is optimized over a continuous range of n.

5.2.4 Bayesian approaches

5.2.4.1 Beta prior for $\boldsymbol{p}$, and discrete uniform prior for $\boldsymbol{n}$

Consider the $\mathrm{Beta}(\alpha, \beta)$ and $\mathrm{Uniform}(1, N)$ priors on p and n respectively (where N is some large preselected integer). These priors have been used by Draper and Guttman [23]. Under the assumption that n and p are independent a priori, and after integrating p out, the marginal posterior of n becomes

$$p(n|X) \propto \frac{\left(kn - \sum_{i=1}^{k} x_i + \beta - 1\right)!}{(kn + \alpha + \beta - 1)!} \prod_{i=1}^{k} \frac{n!}{(n - x_i)!} \quad \text{for } \max x_i \leq n \leq N$$

Draper and Guttman consider the modal value of the marginal posterior distribution as an estimate of n under the absolute loss function.

In the case when $\alpha = \beta = 1$, p follows an arguably noninformative uniform prior distribution. Under that assumption, the authors argue that the choice of the upper bound N does not affect the inferred value of n as long as N does not limit the upper tail. Kahn [170] notes that this is incorrect, since when N increases, both the mean and the median of the posterior of n get arbitrarily large.

Prior information on p can be very easily incorporated into the model by selecting the parameters α and β to match the most likely a priori value of p and the "strength" of the prior belief. Prior information on n is as easy to incorporate, thus the method is nonrestrictive.

5.2.4.2 Uniform prior for $\boldsymbol{p}$, and Poisson prior for $\boldsymbol{n}$

Assume that n follows a Poisson(μ) prior distribution. The observations X_i's then follow a Poisson distribution with mean $\lambda = \mu p$, where λ and p are assumed to be independent a priori. Raftery [169] adopts a hierarchical approach to the problem by specifying a standard vague prior for λ and a uniform prior for p (prior specified in terms of (λ, p) instead of (μ, p)) such that

$$p \sim \text{Uniform}(0,1)$$

$$n \sim \text{Poisson}(\mu) \qquad \text{where } \lambda = \mu p$$

$$p(\lambda) \propto \frac{1}{\lambda}.$$

which leads to a posterior distribution of n, that is

$$p(n|x) \propto \left\{\frac{\left(kn - \sum_{i=1}^{k} x_i\right)!}{(nk+1)!\, n}\right\}\left\{\prod_{i=1}^{k}\binom{n}{x_i}\right\} \qquad \left(n \geq x_{(k)}\right)$$

5.2.4.3 Beta prior for $\boldsymbol{p}$, and Poisson prior for $\boldsymbol{n}$

Hamedani and Walter [171] investigate the case where p follows a beta prior, while n follows a Poisson prior:

$$p \sim \text{Beta}(\alpha, \beta)$$

$$n \sim \text{Poisson}(\lambda).$$

The posterior density for n is:

$$f\left(n|\underset{\sim}{x}\right) = \left(\prod_{i=1}^{k}\binom{n}{x_i}\right)\int_0^1 \frac{\frac{((1-p)^k\lambda)^n}{n!}p^\alpha(1-p)^\beta}{\sum_{m=x_{(k)}}^{\infty}\prod_{i=1}^{k}\binom{m}{x_i}\frac{((1-p)^k\lambda)^m}{m!}}dp\; B(\alpha,\beta) \qquad n \geq x_{(k)}.$$

The resulting estimator for n does not have a simple closed form expression, though.

5.2.4.4 Beta prior for $\boldsymbol{p}$, and left truncated Poisson prior for $\boldsymbol{n}$

Since *n* is the number of trials, *n* has to be greater than or equal to 1. Bayoud [173] assumes that *n* follows a zero truncated Poisson distribution and investigate the model

$$p \sim \text{Beta}(\alpha, \beta)$$

$$n \sim \text{zero truncated Poisson}(\lambda).$$

The marginal posterior pdf of n is then given by

$$h\,(n|x_1, \ldots, x_k) = \frac{1}{C_1}\frac{\lambda^n}{n!}\frac{\Gamma(k\bar{x}+\alpha)\Gamma(kn-k\bar{x}+\beta)}{\Gamma(kn+\alpha+\beta)}\prod_{i=1}^{k}\binom{n}{x_i},$$

where $n \geq \max\{x_{(k)}, 1\}$ and C_1 is the normalizing constant

$$C_1 = \sum\nolimits_{n=\max\{x_{(k)},1\}}^{\infty}\frac{\lambda^n}{n!}\frac{\Gamma(T_k+\alpha)\Gamma(kn-T_k+\beta)}{\Gamma(kn+\alpha+\beta)}\prod_{i=1}^{k}\binom{n}{x_i}.$$

The Bayes point estimate ($\hat{n}_1$) of n under the squared error loss is

$$\hat{n} = \frac{\Gamma(T_k+\alpha)}{C_1}\sum\nolimits_{n=\max\{x_{(k)},1\}}^{\infty} n\frac{\lambda^n}{n!}\frac{\Gamma(kn-T_k+\beta)}{\Gamma(kn+\alpha+\beta)}\prod_{i=1}^{k}\binom{n}{x_i}.$$

Bayoud [173] also discusses how to construct an empirical Bayes estimate for n by obtaining MME values for the hyperparameters α, β and λ.

5.2.4.5 Continuous prior for $\boldsymbol{n}$

Even though the number of trials n is a discrete random variable, as an approximation, Gunel and Chilko [172] considered a continuous prior for n. The model is specified as follows:

$$p \sim \text{Beta}(\alpha, \beta)$$

$$n \sim \text{Gamma}(\alpha + \beta, \delta),$$

where p and n are independent. The justification behind a continuous prior for n stems from the approximated value of n, $n \approx \lambda + \lambda'$, where $\lambda \sim \text{Gamma}(\alpha, \delta)$ and $\lambda' \sim \text{Gamma}(\beta, \delta)$. The estimator does not have a closed form and is evaluated using the Laguerre-Gauss quadrature.

5.2.4.6 Improper prior for $\boldsymbol{n}$

Hamedani and Walter [171] investigate the effect of improper priors for both n and p and show that improper priors lead to implausible results.

5.3 A Beta-Binomial MLE Approach

Let $X_1, \dots, X_k \sim \text{Bin}(n, p)$, where n and p are assumed to be independent, and $n \in \{1,2, \dots\}$ and $p \in (0,1)$. Observe that the likelihood function of n and p given $X_1, \dots, X_k$ is given by:

$$\text{L}(n, p | x_1, \dots, x_k) = \prod_{i=1}^{k} \binom{n}{x_i} p^{x_i} (1-p)^{n-x_i}.$$

The function has to be maximized with respect to p and n simultaneously, where the range for n is restricted to $n \geq \max\{x_{(k)}, 1\}$. Since the parameter p is unknown, we let p fluctuate by viewing it as a random variable arising from the conjugate $\text{Beta}(\alpha, \beta)$ prior distribution given by:

$$g(p|\alpha, \beta) = \frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha, \beta)},$$

where B is the usual beta function, $\alpha, \beta > 0$, and $p \in [0,1]$. The beta distribution lets p take a variety of shapes between 0 and 1, and imposes no severe restrictions on the parameter. After integrating p out, the modified likelihood becomes

$$L(n, \alpha, \beta | x_1, \dots, x_k) = \int \prod_{i=1}^{k} P(x_i|n,p)\, P(p|\alpha,\beta)\, dp$$

$$= \int_0^1 \prod_{i=1}^{k} \binom{n}{x_i} p^{x_i}(1-p)^{n-x_i} \frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha,\beta)}\, dp$$

$$= \left[\prod_{i=1}^{k} \binom{\mathrm{n}}{\mathrm{x_i}}\right] \frac{\mathrm{B}\left(\sum_{i=1}^{k} x_i + \alpha, k\mathrm{n} - \sum_{i=1}^{k} x_i + \beta\right)}{\mathrm{B}(\alpha,\beta)}.$$

A fundamental problem when using the Beta-Binomial distribution is the estimation of the parameters α and β. In this setup, one additional parameter needs to be estimated, n. One possible method is to use the method of moments through the following relation between the first three moments of the Beta-Binomial distribution and the data:

$$E[X_i] = \frac{n\alpha}{\alpha+\beta}$$

$$E[X_i^2] = n\alpha \frac{n\alpha + \beta + n}{(\alpha+\beta)(\alpha+\beta+1)}$$

$$E[X_i^3] = \frac{n\alpha}{(\alpha+\beta)} \frac{\left(b^2 + 3n\beta + 2n^2 + n^2\alpha^2 + \alpha\left(3n^2 + \beta(3n-1)\right)\right)}{(\alpha+\beta+1)(\alpha+\beta+2)}$$

Skellam [181] discussed how to obtain satisfactory estimates of the parameters through the first three moments of the distribution. While this approach is simple and computationally easy, it is not very efficient, and can lead to negative estimates.

A second more efficient, and our preferred, method of obtaining satisfactory estimates of the parameters α, β and n is to use maximum likelihood approach. The log likelihood function

$$l(n, \alpha, \beta) =$$

$$= k \log \Gamma(n+1) - \sum_{i=1}^{k} [\log \Gamma(n - x_i + 1)] + \log \mathrm{B}\left(\sum_{i=1}^{k} x_i + \alpha, kn - \sum_{i=1}^{k} x_i + \beta\right)$$

$$- \log \mathrm{B}(\alpha, \beta)$$

is directly maximized with respect to integer values of $n \geq \max\{x_{(k)}, 1\}$, and continuous values of $\alpha, \beta \geq 0$. Thus, the estimates of n, α, β are given by

$$(\hat{n}, \hat{\alpha}, \hat{\beta}) = \underset{n \geq \max\{x_{(k)}, 1\}, n \in \mathbb{Z}, \alpha > 0, \beta > 0}{\arg\max} l(n, \alpha, \beta).$$

Optimizing the likelihood function with respect to all three parameters, allows the inference based on the data of not only n, but α and β as well. Furthermore, Carroll and Lombard [161] state that if $\alpha, \beta \geq 0$, then the likelihood is maximized at some finite n.

As mentioned earlier, the beta distribution is a flexible distribution with density that can take on a number of shapes. When both $\alpha, \beta < 1$, the density is U-shaped and more sparse. If $\alpha < 1$ or $\beta < 1$, the distribution is reverse J- and J-shaped respectively. When $\alpha = \beta = 1$, the density is flat (uniform) on the unit interval. Finally, as α, β increase, the density "tightens" around its mean and resembles a spike.

To overcome instability and make the $\hat{n}, \hat{\alpha}$, and $\hat{\beta}$ estimators more robust, we restricted the parameter space of α and β to $(0,1000)$, where the upper bound was

somewhat arbitrarily chosen. The imposed restriction is still flexible enough to accommodate all of the above mentioned different shapes of the beta distribution. The lower and upper bounds of α and β permit the mean (variance) of p to vary anywhere from 0 to 1 (0 to 0.25), where p is estimated as $\frac{\alpha}{\alpha+\beta}$. Limiting the upper bound for α and β leads to a more cautious inference about *n*. The final form of our estimator is as follows:

$$\left(\hat{n}, \hat{\alpha}, \hat{\beta}\right) = \underset{n \geq \max\{x_{(k)},1\}, n\in\mathbb{Z}, \alpha,\beta\in(0,1000)}{\arg\max} l(n, \alpha, \beta).$$

We optimized the log likelihood function using a grid search over integer values for n, and a limited memory modification of the BFGS quasi-Newton method proposed by Byrd et al [182]. The BFGS is a quasi-Newton algorithm that was published simultaneously by Broyden, Fletcher, Goldfarb and Shanno Broyden, Fletcher, Goldfarb and Shanno in 1970 [183-186]. We used the *gridSearch* function within the *NMOF* package in R [187, 188], and the *optim* function within the stats package in R [45].

Similar approach has been proposed previously by Carroll and Lombard [161] and Blumenthal and Dahiya [163]. Carroll and Lombard [161] suggested an estimator by maximizing the integrated likelihood as a function of $n \geq x_{(k)}$; α and β were set at some pre-selected integer values. Blumenthal and Dahiya [163] maximized the product $\prod_{i=1}^{k} P(x_i|n,p)\,P(p|\alpha,\beta)$ directly over n and p without integrating p out. They do not provide any guidelines on how to select α and β though.

5.4 Performance Investigation and Applications

5.4.1 Illustrative Example

Consider the random sample $\{16,18,22,25,27\}$ generated from a binomial distribution with $n = 75$ and $p = 0.32$ investigated by Olkin et al [24]. The MME and MLE estimates are 102 and 99 respectively. Suppose that the 27 in the sample was misrecorded, and the correct value was 28. The MME and MLE estimates then become 195 and 190, which shows their lack of robustness. This is an example of an unstable case since the sample mean and sample variance are nearly equal. The proposed estimator based on Beta-Binomial MLE approach is 70 before correcting the sample, and 85 after.

5.4.2 Comparative performance

In this section, we compare the performance of five different $n-$estimators for different combinations of (k, n, p). The five estimators include the stabilized method of moments and stabilized maximum likelihood estimators as presented by Olkin et al Olkin et al [24] (MME:S and MLE:S), the Carroll and Lombard [161] estimators (MB(0,0) and MB(1,1)), and the proposed estimator $\hat{n}$.

First, we present results for eight particularly difficult cases selected by Olkin et al [24]. We also present results for each perturbed sample obtained by adding one to the sample maximum. The results are shown in Table 13. The proposed estimator is reasonably stable when subjected to perturbations. The estimator $\hat{n}$ was closer to the true value of n than the other six estimators in one case. The MME and MLE estimators are highly unstable though. If we disregard those two unstable estimators, $\hat{n}$ was closer/equally close to the

true value of n compared to the other four estimators in two cases. Finally, $\hat{n}$ outperformed both MB(0,0) and MB(1,1) in three cases. It performed similarly well in another three cases, but exhibited higher instability. Sample #6 is of particular interest since it is an unstable case with large p. The MLE:S dominates all estimators followed by the $\hat{n}$ estimator. Finally, if we restrict our attention to sample 8, which is an example of unstable case with small p, we observe that MME:S is the best estimator (excluding MME and MLE due to instability), and the $\hat{n}$ estimator is next.

Table 13: n- estimators for selected samples and perturbed samples.

Sample	n	p	k	MME	MLE	MME:S	MLE:S	MB(0,0)	MB(1,1)	$\hat{n}$
1	75	0.32	5	102 195^{P}	99 191^{P}	70 80^{P}	29 30^{P}	51 57^{P}	49 52^{P}	46 49^{P}
2	34	0.57	4	507 <0 ^{P}	515 ∞^{P}	77 91^{P}	31 32^{P}	52 59^{P}	47 52^{P}	47 55^{P}
3	37	0.17	20	65 154 ^{P}	66 160^{P}	25 27^{P}	11 13^{P}	26 29^{P}	23 25^{P}	31 35^{P}
4	48	0.06	15	18 135 ^{P}	15 127^{P}	10 12^{P}	7 9^{P}	9 12^{P}	8 10^{P}	9 14^{P}
5	40	0.17	12	32 61 ^{P}	40 80^{P}	26 32^{P}	21 22^{P}	27 33^{P}	25 29^{P}	27 42^{P}
6	74	0.68	12	210 259 ^{P}	214 267^{P}	153 162^{P}	67 69^{P}	135 144^{P}	125 131^{P}	109 139^{P}
7	55	0.48	20	71 79 ^{P}	71 81^{P}	69 74^{P}	43 45^{P}	64 70^{P}	63 67^{P}	66 70^{P}
8	60	0.24	15	67 88 ^{P}	67 90^{P}	49 53^{P}	24 26^{P}	45 49^{P}	41 45^{P}	47 51^{P}

Note: The exact samples can be found in Table 2 of Olkin et al [24]
^{P}: perturbed sample

Next, we conducted a simulation study to compare the performance of the five estimators. Following the same design found in the study by Olkin et al [24], we generated values of k, n and p from uniform distributions on $\{3, \dots, 22\}$, $\{1, \dots, 100\}$ and $(0,1)$

respectively. We generated 100,000 random binomial samples, and computed the five estimators for each sample.

The performance of the estimators was evaluated based on three criteria. The first and main criterion is the relative mean squared error, defined for any estimator $\hat{n}$ as $\mathrm{E}\left[\left(\frac{\hat{n}}{n}-1\right)^2\right]$. Several authors including Olkin et al [24], Carroll and Lombard [161], and Casella [162] have suggested that the scaled squared error is an appropriate and natural loss function for this problem. For easier interpretation, we reported relative mean square error efficiency instead. A relative mean square error efficiency of an estimator $\hat{n}$ relative to a benchmark estimator $\bar{n}$ is the ratio of the relative mean square error of $\bar{n}$ to the relative mean square error of $\hat{n}$. The second criterion is the bias, defined as $\mathrm{E}\left[\frac{\hat{n}}{n}-1\right]$ similarly to the study of Blumenthal and Dahiya [163]. Lastly, we computed the number of times each estimator "won" (was closest to the true value of n). Ties were counted as wins.

Finally, each sample was categorized as either stable or unstable. We used the criterion that Olkin et al [24] suggested – a sample is classified as stable if and only if $\frac{\bar{x}}{s^2} \geq 1+\frac{1}{\sqrt{2}}$.

The relative mean square errors efficiency, bias and number of wins for both the stable and unstable cases are showed in Table 14.

Overall, the $\hat{n}$ estimator performed slightly better than the other four estimators with an overall efficiency gain of about 4% and 1% over MME:S and MB(1,1) respectively. In terms of proximity to the true n, the $\hat{n}$ estimator won/tied in 42% of the cases, while the

MB(1,1) and MME:S both won/tied in about 50% of the cases. The MLE:S estimator performed the worst in terms of efficiency and bias, and second worst behind MB(0,0) in terms of number of wins. The MB(1,1) outperformed MB(0,0) across all three measures of error.

However, in stable cases, which occurred 69.4% of the time, the $\hat{n}$ estimator performed slightly worse than or equally well at best compared to the other estimators with an overall efficiency loss of 1% over MME:S. MB(1,1) had an overall efficiency gain of 1% over MME:S and won/tied in 63% of the cases. The MB(1,1) performed the best across all three measures of error. MB(1,1) again outperformed MB(0,0) across all three measures of error.

In unstable cases (30.6% of all cases), the $\hat{n}$ estimator dominated all other estimators, and showed an 8% overall efficiency gain over MME:S, and about 4% over both MB(0,0) and MB(1,1). In terms of bias and number of wins, MME:S performed the best followed by the proposed $\hat{n}$ estimator. This time, MB(1,1) performed similarly to or worse than MB(0,0) across all three measures of error. The $\hat{n}$ estimator outperformed MLE:S, MB(0,0) and MB(1,1) across every criterion.

Table 14: Comparison of the n-estimators.

	Cases		MME:S	MLE:S	MB(0,0)	MB(1,1)	$\hat{n}$
All cases	10000	Efficiency	1.00	0.93	1.02	1.03	1.04
		Bias	-0.186	-0.304	-0.223	-0.23	-0.218
		Wins	49%	35%	33%	50%	42%
Stable	6939	Efficiency	1.00	0.99	0.99	1.01	0.99
	(69.4%)	Bias	-0.172	-0.179	-0.185	-0.165	-0.187
		Wins	49%	42%	39%	64%	42%
Unstable	3061	Efficiency	1.00	0.9	1.05	1.05	1.09
	(30.6%)	Bias	-0.217	-0.583	-0.309	-0.375	-0.286
		Wins	47%	19%	19%	19%	41%

We investigated the performance of the estimators further by splitting the parameter range for p into several overlapping categories following Olkin et al [24]. For stable cases, MB(1,1) achieved highest efficiency across values for $p \in (0.2,1)$, which can be explained by the tendency of the MB(0,0) and MB(1,1) estimators to downweigh the possibility that p is near zero. For "small" values of p, $\left(0 < p < \sqrt{2} - 1\right)$, the MME:S achieved the best performance.

In the special unstable case with "small p" $(0 < p < \sqrt{2} - 1)$ the proposed $\hat{n}$ estimator was superior to the other estimators, followed by MME:S. For bigger values of p the MB(1,1) performed better than all other estimators.

5.5 n −estimator when $k = 1$ with Applications in Contingency Tables

5.5.1 Background

In the absence of replications, inference about n is not possible. And yet, the need for an n −estimator, even when only a single observation is available, arises in certain

situations, particularly in analyzing partially reported contingency tables when the interest lies in inference about unobserved cell counts.

Consider a simple 2×2 contingency table design, where the columns are counts of subjects who have a particular disease of interest and those who do not, and the rows are counts of those subjects who tested positive and negative for the disease. A standard assumption is that the disease and no disease groups are independent. Suppose that the test found agreement in A subjects for being positive (true positives), and in D subjects for being negative (true negatives). Let B represent the number of subjects who do not have the disease and tested positive (false positives), and C the number of subjects who have the disease and tested negative (false negatives). The contingency table design for evaluating the performance of a test can be represented as Table 15.

Table 15: Contingency design table

	Disease	No disease	Total
Test positive	A	B	$A + B$
Test negative	C	D	$C + D$
Total	$A + C$	$B + D$	$A + B + C + D$

Some studies report only the number of true positives and true negatives, that is cell counts A and D. When B and C are not reported, the total number of subjects who have the disease $A + C$ and the total number of subjects who do not have the disease $B + D$ are unknown. This is problematic since in order to determine a test's performance, a key quantity of interest is the sensitivity and specificity of the test. Sensitivity is the ability of

a test to identify correctly those who have the disease, and specificity is the ability to identify correctly those who do not have the disease. The sensitivity (true positive rate) and specificity (true negative rate) of a test are defined as:

$$\text{Sensitivity} = \frac{\mathrm{A}}{\mathrm{A} + \mathrm{C}^*}$$

$$\text{Specificity} = \frac{D}{B^* + D},$$

where * refers to counts that are not reported.

With the increase in knowledge and technology in the medical field in recent years, an increasing number of diagnostic tests for different diseases has become available. Diagnostics tests should be used based on their validity, and not availability. Validity of a test can be characterized by its sensitivity and specificity.

This section addresses the problem of inference on the sensitivity and specificity for a test in a missing data context when only the true positives A and true negatives D are reported. Equivalently, the problem can be translated in the context of the binomial n problem with $k = 1$ replications. The problem now is to estimate column totals $\mathrm{A} + \mathrm{C}^*$ and $B^* + D$ conditional on the observed values of some cell counts (true positives A and true negatives D). We propose a Bayesian model to recover the contingency table by incorporating any prior information available.

5.5.2 Bayesian Model

Assume that Table 15 is partially observed, that is A and D are known, while B and C are unknown. Since the disease and no disease populations are independent from one another, we can treat each of A and D as a realization of an independent binomial distribution as follows:

$$A \sim \text{Binomial}(n_1, p_1)$$

$$D \sim \text{Binomial}(n_2, p_2),$$

where $n_1 = A + C^*$, $n_2 = B^* + D$, and the probabilities p_1 and p_2 of testing positive (negative) in the presence (absence) of the disease are unknown. A natural restriction is that $n_1 \geq A$ and $n_2 \geq D$.

We propose a Bayesian approach to the problem. We assume that the sample sizes n_1 and n_2 follow a Poisson distribution with parameters λ_1 and λ_2 respectively. The Poisson distribution accommodates the discrete nature of the parameters. Additionally, we let the probabilities p_1 and p_2 each be a draw from a uniform distribution

Since sensitivity and specificity are both usually in the interval $[0.5,1]$, we let p_1 be a draw from a Uniform$(0.5,1)$ distribution.

In practice, for any test, there is a tradeoff between sensitivity and specificity. Due to a minimum error bound, as sensitivity increases, specificity decreases, and vice versa. Therefore, to quantify and incorporate this tradeoff into our model, we let sensitivity and

specificity together ($p_1 + p_2$) vary from 1.4 to 1.8. Therefore, we assume that p_2 follows a Uniform$(1.4 - p_1, 1.8 - p_1)$ distribution.

Finally, λ_1 and λ_2 are each selected based on prior knowledge about $A + C^*$ and $B^* + D$ respectively. The $[0.5,1]$ range for sensitivity and specificity limits the range for n_1 and n_2 to $[A, 2A]$ and $[D, 2D]$ respectively. Using this restriction, if there is no prior information available, we assign the parameters uninformative priors by centering λ_1 and λ_2 at $\frac{4}{3}A$ and $\frac{4}{3}D$.

The model has the following form:

$$A|n_1, p_1 \sim \text{Binom}(n_1, p_1) \qquad D|n_2, p_2 \sim \text{Binom}(n_2, p_2)$$

$$p_1 \sim \text{Uniform}(0.5,1) \qquad p_2|p_1 \sim \text{Uniform}(1.4 - p_1, 1.8 - p_1)$$

$$n_1|\lambda_1 \sim \text{Poi}(\lambda_1) \qquad n_2|\lambda_2 \sim \text{Poi}(\lambda_2)$$

We generated samples from the posterior distribution using Markov chain Monte Carlo (MCMC) algorithm implemented in JAGS via the *rjags* package in R. We used 1 chain with the first 10,000 iterations discarded while the Markov chain stabilized. Assuming quadratic loss function, posterior inference was based on posterior means generated from 50,000 samples thinned at a lag of 50.

5.5.3 Example

In this section, we demonstrate how our methodology can be used to recover contingency tables and to estimate sensitivity and specificity. Consider the following fully

observed contingency table from a study by Tubman et al [189] . The study investigated the performance of different screening tests for detecting congenital heart disease early in the life of children with Down's syndrome. Table 16 shows the diagnostic ability of a combination of clinical examination, chest radiography and electrocardiography:

Table 16: Original contingency table reported by Tubman et al [183]

	Disease	No disease
Test positive	24	4
Test negative	10	43

Suppose that only the true positive (24) and true negative (43) diagnoses of heart disease are reported. Without incorporating any additional information about the association between the cell counts, the resulting recovered contingency table based on the posterior means obtained through our methodology is presented in Table 17.

Table 17: Recovered contingency table with uninformative prior

	Disease	No disease
Test positive	24	13
Test negative	7	43

Considering that the model is based on a single observation and due to the prior being fairly non-informative, the performance of the model is not remarkable. The estimated false negatives were 7 (compared to the true value 10), and the estimated false positives were 13

(compared to the true value 4). While the estimated sensitivity of 0.77 is relatively close to the actual sensitivity of the test, 0.71, it is apparent that the estimated specificity of 0.77 is significantly lower than the true value, 0.91.

Incorporating more information into the model through the priors leads to a significantly better performance. Figure 21 and Figure 22 below show the absolute deviation of the recovered cell counts B and C from their true values as we vary n_1 and n_2 through the parameters λ_1 and λ_2. We let n_1 vary anywhere from A to $2A$ (that is from 24 to 48). Similarly, we let n_2 vary anywhere from D to $2D$ (that is from 43 to 86). The true values for n_1 and n_2 are 34 and 47 respectively.

Unsurprisingly, Figure 21 shows that the deviation of the estimate of B from the original table increases as we increase λ_2 which governs the distribution on n_2 ($n_2 = B + D$). For any prior of λ_2 between 43 and 45, the posterior of B (n_2) stayed within 3 deviations from the original cell count of B, 4 (original count n_2,47). As λ_2 increased, the error increased. The estimator of B was not affected by the prior on λ_1 which governs the distribution of n_1.

The effect of the priors on the posterior of C showed an interesting pattern that can be observed in Figure 22. The two priors, λ_1 and λ_2 both influence C, and the posterior estimate of C does not remain the same as we vary the prior on λ_2. The posterior of C generated by the model was within 1 deviation from the original value of 10 12.5% of the scenarios, and within 3 deviations from the original value about 37.49% of the time.

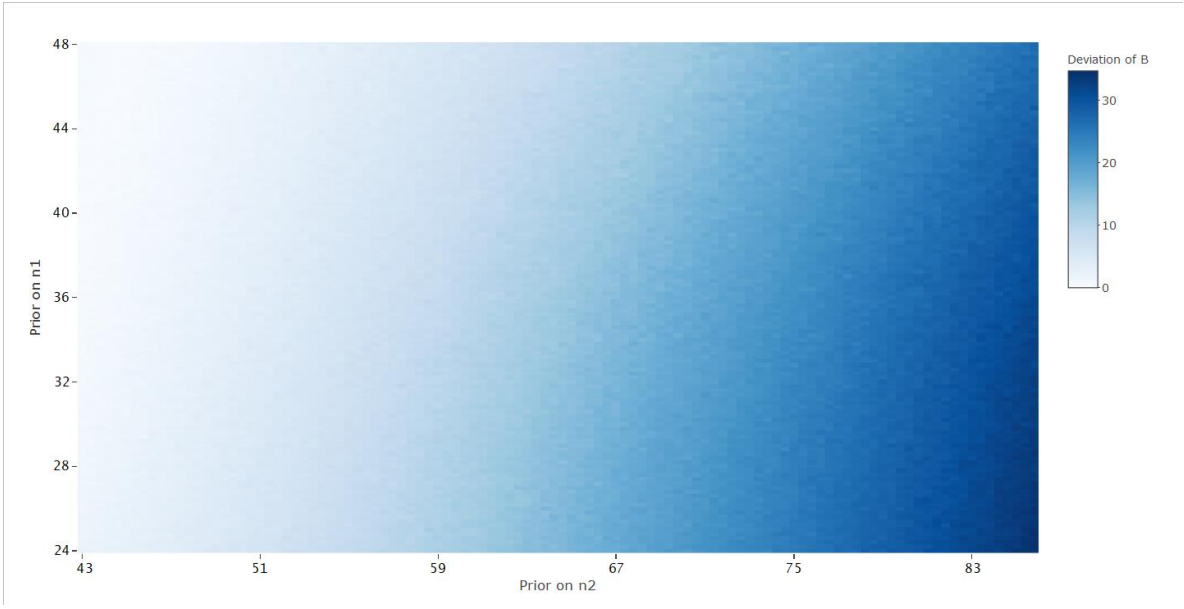

Figure 21. Deviation of B

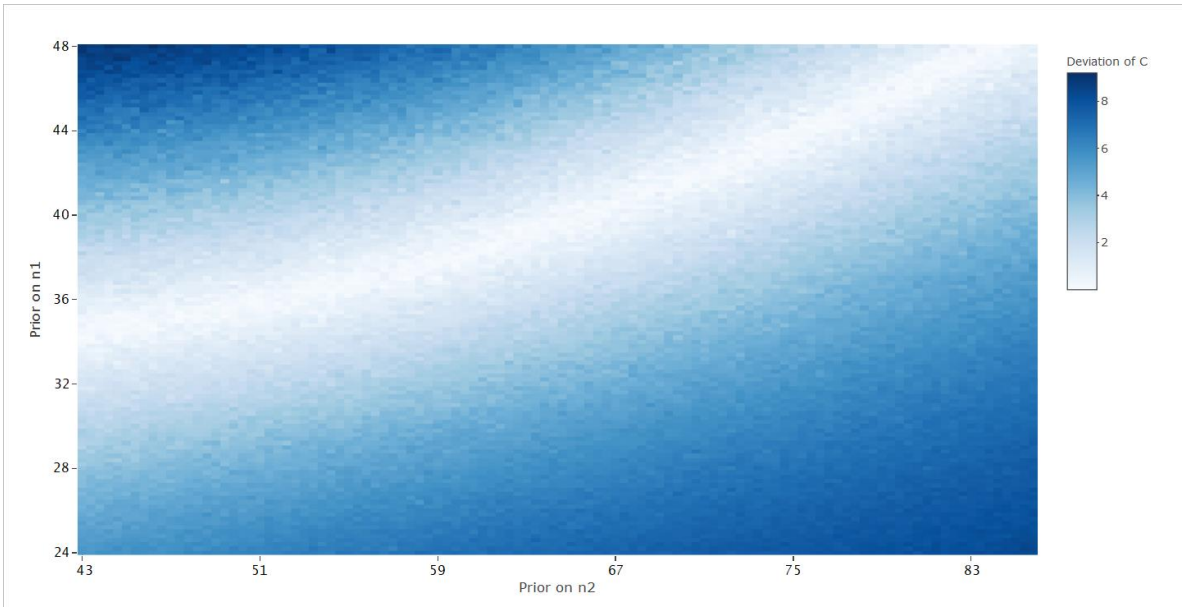

Figure 22. Deviation of C

APPENDIX A. SUPPLEMENTAL MATERIAL FOR CHAPTER 2

A.1 Full search strategy for PubMed

Detailed search terms and strategy for title/abstract: (melanom* or melanocyt*) AND (BRAF or BRAF* or *RAF or MEK1 or MEK2 or MEK* or MAPK or ERK1 or ERK2 or ERK* or R05185426 or RG7204 or PLX 4032 or vemurafenib or zelboraf or dabrafenib or tanfilar or GSK 2118436 or GSK2118436 or GSK-2118436 or JTP 74057 or trametinib or mekinist or JTP74057 or JTP-74057 or GSK1120212 or GSK1120212 or GSK-1120212 or cobimetinib or cotellic or GDC-0973 or XL518 or pd-1 or pd-l1 or pd-l2 or programmed cell death receptor or programmed cell death 1 receptor or programmed cell death 2 receptor or CD279 or CLTA-4 or Cytotoxic T-Lymphocyte-Associated Antigen 4 or Cytotoxic T Lymphocyte Associated Antigen 4 or Cytotoxic T-Lymphocyte Antigen 4 or Cytotoxic T Lymphocyte Antigen 4 or CD152 or CD28 OR ipilimumab or MDX-CTLA-4 or Yervoy or MDX 010 or MDX010 or MDX-010 or tremelimumab or ticilimumab or CP 675 or CP675 or CP-675 or CP-675,206 or CP-675206 or CP675206 or CP 675206 or pidilizumab or nivolumab or opdivo or bms-936558 or ono-4538 or ono4538 or mdx-1106 or pembrolizumab or lambrolizumab or keytruda or mk-3475) AND (random* or randomised or randomized or prospective).

A.2 Cochrane risk of bias tool

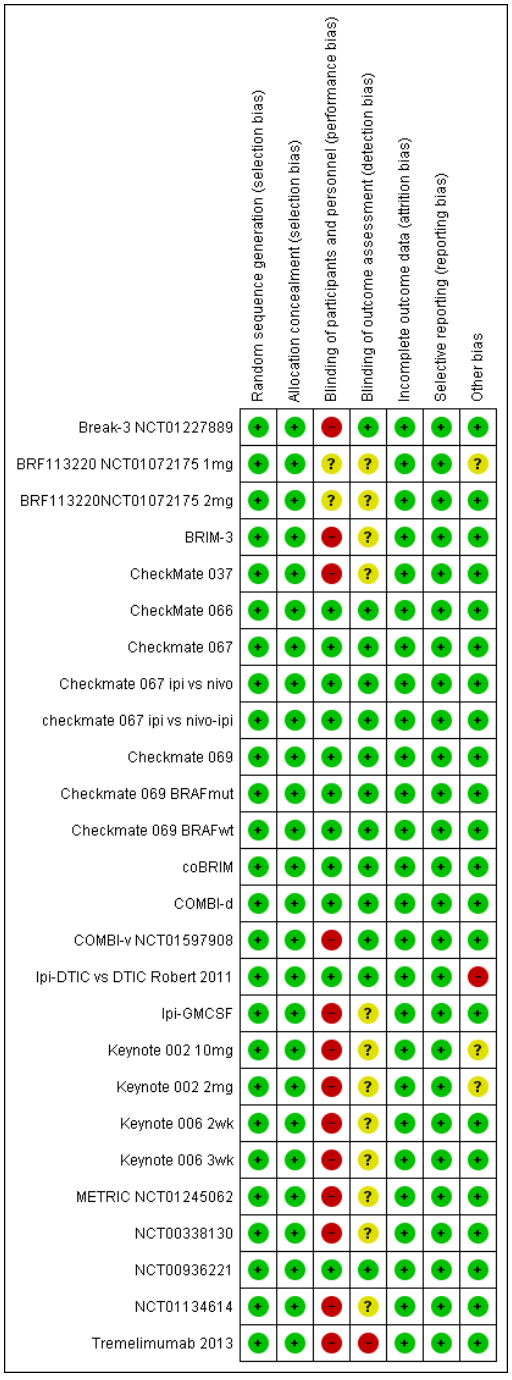

Figure 23. Risk of bias analysis

A.3 Funnel plot of publication bias

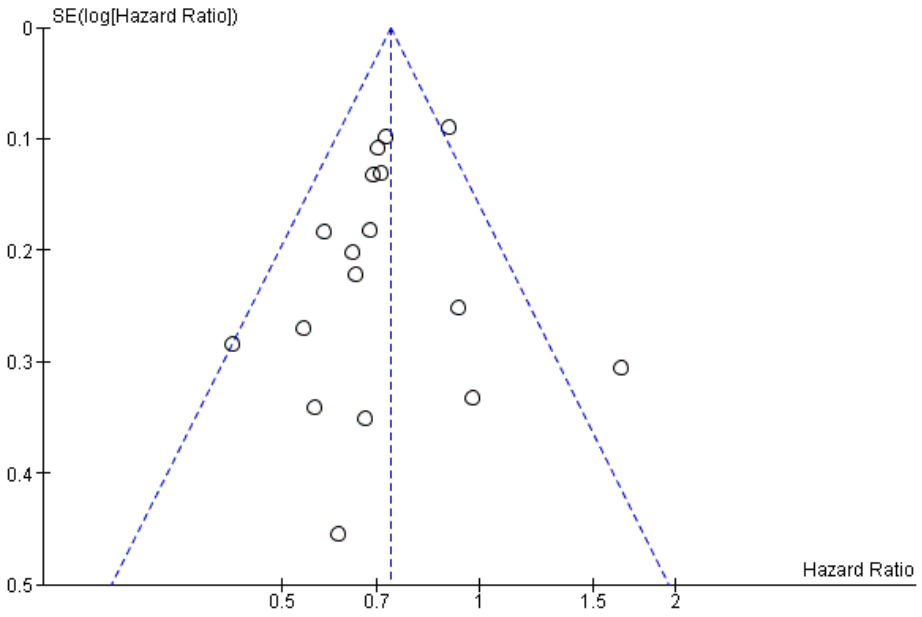

Figure 24. Funnel plot of all included studies (Overall survival outcome)

APPENDIX B. SUPPLEMENTAL MATERIAL FOR CHAPTER 3

Table 18: All reported adverse events for enzalutamide vs. abiraterone in the pre- and post-docetaxel settings

		Enzalutamide, pre-docetaxel (PREVAIL)	Placebo, pre-docetaxel (PREVAIL)	Abiraterone, pre-docetaxel (COU-AA-302)	Placebo/Prednisone, pre-docetaxel (COU-AA-302)	Enzalutamide, post-docetaxel (AFFIRM)	Placebo, post-docetaxel (AFFIRM)	Abiraterone, post-docetaxel (COU-AA-301)	Placebo/Prednisone, post-docetaxel (COU-AA-301)
Adverse Event	Any grade Grade ≥3	0.97 (0.96-0.98) 0.46 (0.42-0.49)	0.93 (0.91-0.95) 0.37 (0.34-0.41)	0.99 (0.98-1.00) 0.48 (0.43-0.52)	0.97 (0.95-0.98) 0.42 (0.38-0.46)	0.98 (0.97-0.99) 0.45 (0.42-0.49)	0.98 (0.96-0.99) 0.53 (0.48-0.58)	0.77 (0.74-0.80) 0.23 (0.20-0.26)	0.77 (0.73-0.81) 0.19 (0.16-0.23)
Any SAE		0.36 (0.33-0.39)	0.27 (0.24-0.30)	0.33 (0.29-0.37)	0.26 (0.23-0.30)	0.34 (0.30-0.37)	0.39 (0.34-0.43)		
AE leading to treatment discontinuation		0.06 (0.04-0.07)	0.06 (0.05-0.08)	0.10 (0.08-0.13)	0.09 (0.07-0.12)	0.08 (0.06-0.10)	0.10 (0.07-0.13)	0.13 (0.11-0.16)	0.18 (0.15-0.22)
AE leading to death		0.04 (0.03-0.06)	0.04 (0.03-0.05)	0.04 (0.02-0.06)	0.02 (0.01-0.04)	0.03 (0.02-0.04)	0.04 (0.02-0.06)	0.13 (0.11-0.16)	0.15 (0.12-0.19)
Any cardiac	Any grade Grade ≥3	0.10 (0.08-0.12) 0.03 (0.02-0.04)	0.08 (0.06-0.10) 0.02 (0.01-0.03)	0.19 (0.16-0.22) 0.06 (0.04-0.08)	0.16 (0.13-0.19) 0.03 (0.02-0.05)	0.06 (0.05-0.08) 0.01 (0.00-0.02)	0.08 (0.05-0.11) 0.02 (0.01-0.04)	0.16 (0.14-0.19) 0.05 (0.04-0.07)	0.12 (0.09-0.15) 0.02 (0.01-0.04)
Cardiac AE leading to death								0.01 (0.01-0.02)	0.01 (0.01-0.03)
Atrial fibrillation	Any grade Grade ≥3	0.02 (0.01-0.03) 0.00 (0.00-0.01)	0.01 (0.01-0.02) 0.01 (0.00-0.01)	0.04 (0.03-0.06) 0.01 (0.01-0.03)	0.05 (0.03-0.07) 0.01 (0.00-0.02)				
Acute coronary syndromes/myocardial infarction	Any grade Grade ≥3	0.01 (0.00-0.02) 0.01 (0.00-0.02)	0.00 (0.00-0.01) 0.00 (0.00-0.01)			0.00 (0.00-0.01) 0.00 (0.00-0.01)	0.01 (0.00-0.02) 0.01 (0.00-0.02)		
Acute renal failure	Any grade Grade ≥3	0.04 (0.03-0.05) 0.01 (0.01-0.02)	0.05 (0.03-0.06) 0.01 (0.01-0.02)						

Table 18 (Continued)

Ischemic or hemorrhagic cerebrovascular event	Any grade Grade ≥3	0.01 (0.01-0.02) 0.01 (0.00-0.01)	0.01 (0.01-0.02) 0.00 (0.00-0.01)						
ALT elevation	Any grade Grade ≥3	0.01 (0.00-0.02) 0.00 (0.00-0.01)	0.01 (0.00-0.01) 0.00 (0.00-0.01)	0.12 (0.09-0.15) 0.05 (0.04-0.08)	0.05 (0.03-0.07) 0.01 (0.00-0.02)				
AST elevation	Any grade Grade ≥3			0.11 (0.08-0.14) 0.03 (0.02-0.05)	0.05 (0.03-0.07) 0.01 (0.00-0.02)				
Abnormalities in liver function tests	Any grade Grade ≥3					0.01 (0.01-0.02) 0.00 (0.00-0.01)	0.02 (0.01-0.03) 0.01 (0.00-0.02)	0.11 (0.09-0.14) 0.04 (0.03-0.05)	0.09 (0.06-0.12) 0.04 (0.02-0.06)
Seizure	Any grade Grade ≥3	0.00 (0.00-0.01) 0.00 (0.00-0.01)	0.00 (0.00-0.01) 0.00 (0.00-0.00)			0.01 (0.00-0.01) 0.01 (0.00-0.01)	0.00 (0.00-0.01) 0.00 (0.00-0.01)		
Anemia	Any grade Grade ≥3							0.25 (0.22-0.28) 0.08 (0.06-0.10)	0.28 (0.24-0.33) 0.08 (0.06-0.11)
Thrombocytopenia	Any grade Grade ≥3							0.04 (0.03-0.05) 0.01 (0.01-0.02)	0.04 (0.02-0.06) 0.01 (0.00-0.02)
Neutropenia	Any grade Grade ≥3							0.01 (0.01-0.02) 0.00 (0.00-0.01)	0.01 (0.00-0.02) 0.00 (0.00-0.01)
Febrile neutropenia	Any grade Grade ≥3							0.00 (0.00-0.01) 0.00 (0.00-0.01)	0.00 (0.00-0.01) 0.00 (0.00-0.01)
Abdominal pain	Any grade Grade ≥3							0.13 (0.11-0.15) 0.02 (0.01-0.04)	0.12 (0.09-0.16) 0.02 (0.01-0.04)

Table 18 (Continued)

Arthralgia	Any grade Grade ≥3	0.21 (0.19-0.24) 0.02 (0.01-0.03)	0.16 (0.14-0.19) 0.01 (0.01-0.02)	0.28 (0.25-0.32)	0.24 (0.20-0.28)			0.30 (0.27-0.34) 0.05 (0.04-0.07)	0.24 (0.20-0.29) 0.04 (0.03-0.07)
Asthenia	Any grade Grade ≥3	0.47 (0.44-0.50) 0.03 (0.02-0.05)	0.33 (0.30-0.36) 0.03 (0.02-0.04)					0.15 (0.13-0.18) 0.03 (0.02-0.05)	0.14 (0.11-0.17) 0.02 (0.01-0.04)
Back pain	Any grade Grade ≥3	0.29 (0.26-0.32) 0.03 (0.02-0.04)	0.22 (0.20-0.25) 0.03 (0.02-0.04)	0.32 (0.28-0.36)	0.32 (0.28-0.36)			0.33 (0.30-0.36) 0.07 (0.05-0.09)	0.36 (0.31-0.41) 0.10 (0.08-0.14)
Bone pain	Any grade Grade ≥3			0.20 (0.16-0.23)	0.19 (0.16-0.23)			0.27 (0.24-0.31) 0.06 (0.05-0.08)	0.30 (0.25-0.34) 0.08 (0.06-0.11)
Constipation	Any grade Grade ≥3	0.23 (0.21-0.26) 0.01 (0.00-0.01)	0.17 (0.15-0.20) 0.00 (0.00-0.01)	0.23 (0.20-0.27)	0.19 (0.16-0.23)			0.28 (0.25-0.31) 0.01 (0.01-0.02)	0.32 (0.28-0.37) 0.01 (0.00-0.03)
Cough				0.17 (0.14-0.21)	0.14 (0.11-0.17)				
Decreased appetite	Any grade Grade ≥3	0.19 (0.16-0.22) 0.00 (0.00-0.01)	0.16 (0.14-0.19) 0.01 (0.00-0.02)						
Diarrhea	Any grade Grade ≥3	0.17 (0.14-0.19) 0.00 (0.00-0.01)	0.14 (0.12-0.17) 0.00 (0.00-0.01)	0.22 (0.18-0.25)	0.18 (0.15-0.21)	0.21 (0.19-0.24) 0.01 (0.01-0.02)	0.18 (0.14-0.22) 0.00 (0.00-0.01)	0.20 (0.17-0.23) 0.01 (0.01-0.02)	0.15 (0.12-0.19) 0.01 (0.01-0.03)
Dizziness	Any grade Grade ≥3	0.11 (0.09-0.14) 0.00 (0.00-0.01)	0.07 (0.06-0.09) 0.00 (0.00-0.00)						
Dysgeusia	Any grade Grade ≥3	0.08 (0.06-0.10) 0.00 (0.00-0.01)	0.04 (0.03-0.05) 0.00 (0.00-0.00)						
Peripheral edema/edema	Any grade Grade ≥3	0.11 (0.10-0.14) 0.00 (0.00-0.01)	0.08 (0.07-0.10) 0.00 (0.00-0.01)	0.28 (0.24-0.32) 0.01 (0.00-0.02)	0.24 (0.20-0.27) 0.02 (0.01-0.03)			0.33 (0.30-0.36) 0.03 (0.02-0.04)	0.24 (0.20-0.28) 0.01 (0.00-0.03)

Table 18 (Continued)

Fall	Any grade Grade ≥3	0.13 (0.11-0.15) 0.02 (0.01-0.03)	0.05 (0.04-0.07) 0.01 (0.00-0.02)						
Dyspnea	Any grade Grade ≥3	0.11 (0.09-0.13) 0.01 (0.00-0.01)	0.09 (0.07-0.11) 0.01 (0.00-0.01)					0.15 (0.12-0.17) 0.02 (0.01-0.03)	0.12 (0.10-0.16) 0.02 (0.01-0.04)
Fatigue	Any grade Grade ≥3			0.39 (0.35-0.43)	0.34 (0.30-0.38)	0.34 (0.30-0.37) 0.06 (0.05-0.08)	0.29 (0.25-0.34) 0.07 (0.05-0.10)	0.47 (0.44-0.51) 0.09 (0.07-0.11)	0.44 (0.39-0.49) 0.10 (0.08-0.14)
Gynecomastia	Any grade Grade ≥3	0.03 (0.02-0.05) 0.00 (0.00-0.00)	0.01 (0.01-0.02) 0.00 (0.00-0.00)						
Headache	Any grade Grade ≥3	0.11 (0.09-0.13) 0.00 (0.00-0.01)	0.07 (0.05-0.09) 0.00 (0.00-0.01)			0.12 (0.10-0.14) 0.01 (0.00-0.02)	0.06 (0.04-0.08) 0.00 (0.00-0.01)		
Hematuria	Any grade Grade ≥3	0.09 (0.07-0.11) 0.01 (0.01-0.02)	0.06 (0.04-0.08) 0.01 (0.01-0.02)					0.09 (0.07-0.11) 0.02 (0.01-0.03)	0.09 (0.06-0.12) 0.02 (0.01-0.04)
Hot flush	Any grade Grade ≥3	0.18 (0.16-0.21) 0.00 (0.00-0.01)	0.08 (0.06-0.10) 0.00 (0.00-0.00)	0.22 (0.19-0.26)	0.18 (0.15-0.22)	0.20 (0.18-0.23) 0.00 (0.00-0.00)	0.10 (0.08-0.14) 0.00 (0.00-0.01)		
Hypertension	Any grade Grade ≥3	0.14 (0.12-0.17) 0.07 (0.06-0.09)	0.04 (0.03-0.06) 0.02 (0.01-0.03)	0.22 (0.19-0.25) 0.04 (0.03-0.06)	0.13 (0.11-0.16) 0.03 (0.02-0.05)	0.07 (0.05-0.09)	0.03 (0.02-0.05)	0.11 (0.09-0.14) 0.01 (0.01-0.02)	0.08 (0.06-0.11) 0.00 (0.00-0.01)
Hypokalaemia	Any grade Grade ≥3			0.17 (0.14-0.20) 0.02 (0.01-0.04)	0.13 (0.10-0.16) 0.02 (0.01-0.03)			0.18 (0.16-0.21) 0.04 (0.03-0.06)	0.09 (0.07-0.12) 0.01 (0.00-0.02)
Insomnia	Any grade Grade ≥3	0.08 (0.07-0.10) 0.00 (0.00-0.01)	0.06 (0.04-0.07) 0.00 (0.00-0.00)						

Table 18 (Continued)

Lower respiratory tract or lung infection	Any grade Grade ≥3	0.08 (0.06-0.10) 0.01 (0.01-0.03)	0.05 (0.03-0.06) 0.01 (0.01-0.02)						
Upper respiratory tract infection	Any grade Grade ≥3	0.16 (0.14-0.19) 0.00 (0.00-0.00)	0.11 (0.09-0.13) 0.00 (0.00-0.00)						
Mental impairment disorders	Any grade Grade ≥3	0.06 (0.04-0.07) 0.00 (0.00-0.00)	0.01 (0.01-0.02) 0.00 (0.00-0.01)						
Muscle spasm				0.14 (0.11-0.17)	0.20 (0.17-0.24)				
Musculoskeletal pain	Any grade Grade ≥3					0.14 (0.11-0.16) 0.01 (0.01-0.02)	0.10 (0.07-0.13) 0.00 (0.00-0.01)		
Nausea	Any grade Grade ≥3			0.22 (0.19-0.26)	0.22 (0.19-0.26)			0.33 (0.29-0.36) 0.02 (0.01-0.03)	0.33 (0.29-0.38) 0.03 (0.02-0.05)
Non-pathological fracture	Any grade Grade ≥3	0.09 (0.07-0.11) 0.02 (0.01-0.03)	0.03 (0.02-0.04) 0.01 (0.01-0.02)						
Pain	Any grade Grade ≥3							0.05 (0.04-0.07) 0.01 (0.00-0.02)	0.05 (0.04-0.08) 0.02 (0.01-0.04)
Pain in extremity	Any grade Grade ≥3			0.17 (0.14-0.20)	0.16 (0.13-0.19)			0.20 (0.17-0.23) 0.03 (0.02-0.04)	0.21 (0.17-0.25) 0.05 (0.03-0.08)
Pyrexia	Any grade Grade ≥3							0.10 (0.08-0.12) 0.00 (0.00-0.01)	0.09 (0.07-0.12) 0.01 (0.01-0.03)
Restless leg syndrome	Any grade Grade ≥3	0.02 (0.01-0.03) 0.00 (0.00-0.01)	0.00 (0.00-0.01) 0.00 (0.00-0.00)						

Table 18 (Continued)

Urinary tract infection	Any grade Grade ≥3							0.13 (0.11-0.16) 0.02 (0.01-0.03)	0.07 (0.05-0.10) 0.01 (0.00-0.02)
Vomiting	Any grade Grade ≥3							0.24 (0.21-0.27) 0.03 (0.02-0.04)	0.26 (0.22-0.30) 0.03 (0.02-0.05)
Weight decrease	Any grade Grade ≥3	0.12 (0.10-0.15) 0.01 (0.00-0.02)	0.09 (0.07-0.11) 0.00 (0.00-0.01)						

Table 19: Sensitivity analysis meta-estimates for enzalutamide vs. abiraterone in the pre- and post-docetaxel settings

		Overall Survival	Radiographic Progression-Free Survival	Time to PSA Progression	PSA Response Rate
Enzalutamide vs. Abiraterone, Pre-Docetaxel	Posterior Median Hazard Ratio	0.91 (95% CrI 0.45-1.84, 95% PrI 0.35-2.37)	0.61 (95% CrI 0.30-1.25, 95% PrI 0.24-1.59)	0.36 (95% CrI 0.18-0.75, 95% PrI 0.14-0.96)	6.01 (95% CrI 2.30-15.82, 95% PrI 1.88-19.33)[b]
	Posterior Probability Hazard Ratio < 1[a]	0.60 (0.58)	0.91 (0.85)	1.00 (0.98)	1.00 (1.00)[b]
Enzalutamide vs. Abiraterone, Post-Docetaxel	Posterior Median Hazard Ratio	0.90 (95% CrI 0.44-1.83, 95% PrI 0.35-2.34)	0.61 (95% CrI 0.30-1.26, 95% PrI 0.24-1.60)	0.38 (95% CrI 0.18-0.78, 95% PrI 0.14-0.99)	7.17 (95% CrI 2.50-20.60, 95% PrI 2.08-24.83)[b]
	Posterior Probability Hazard Ratio < 1[a]	0.62 (0.59)	0.91 (0.85)	0.99 (0.98)	1.00 (1.00)[b]

[a]Posterior probability (predictive probability); [b]odds ratio for response.

Table 20: Characteristics of studies included in the sequencing assessment

Study Name	Patient population (N)	Median follow up (month)	Prior Treatment (n)	Study treatment	Median Survival (months)	
					OS (95%CI)	PFS (95%CI)
Ryan et al.	Metastatic CRPC pre-docetaxel (N=1088)	49.2	Docetaxel	Abiraterone (n=546)	34.7 (32.7-36.8)	33.4 (30.2-39.8)
				Placebo (n=542)	30.3 (28.7-33.3)	23.4 (20.3-27.5)
Rathkoph et al.	Metastatic CRPC pre-docetaxel (N=1088)	27.1	Docetaxel	Abiraterone (n=546)	34.7 (32.7-36.8)*	13.5 (10.9-14.8)*
				Placebo (n=542)	30 (28.7-33.3)*	8.2 (8.1-8.5)*
Scher et al.	Progressive CRPC post-docetaxel (N=1199)	14.4	Docetaxel	Enzalutamide (n=800)	18.4 (17.3-NR)	8.3 (8.2-9.4)
				Placebo (n=399)	13.6 (11.3-15.8)	2.9 (2.8-3.4)
Loriot et al.	Metastatic CRPC (N=38)	5.8*	Docetaxel, Enzalutamide	Abiraterone (n=38)	7.2 (5-NR)	2.7 (2.3-4.1)
Noonan et al.	Metastatic CRPC (N=27)	6.8*	Docetaxel, Enzalutamide	Abiraterone (n=27)	11.5 (6.5.3-16.6)	3.5 (2.5-4.6)
Badrising et al.	Metastatic CRPC (N=61)	3.8	Docetaxel, Abiraterone	Enzalutamide (n=61)	7.3 (6.6-NR)	2.8 (2.6-3.7)
Bianchini et al.	(N=39)	4.3	Docetaxel, Abiraterone	Enzalutamide (n=39)	NR	2.8 (2-3.6)
Brasso et al.	(N=137)	6.5*	Docetaxel, Abiraterone	Enzalutamide (n=137)	8.3 (6.8-9.8)	-

Table 20 (Continued)

Schmid et al.	(N=35)	5	Docetaxel, Abiraterone	Enzalutamide (n=35)	7.5 (4.7-10.3)	3.1 (1.4-4.8)
Schrader et al.	(N=35)	5.3*	Docetaxel, Abiraterone	Enzalutamide (n=35)	7.1 (6.2-8.1)	4 (2-6)
Cheng et al.	(N=165)	6.3*	Docetaxel, Abiraterone	Enzalutamide (n=165)	12.2 (10.7-16.5)	2.8 (2.5-3.2)

*Estimated parameters (not reported)
NR, not reached

APPENDIX C. SUPPLEMENTAL MATERIAL FOR CHAPTER 4

C.1 Heterogeneity

Table 21: Posterior median (95% CrI) costs (2018 USD), QALYs and ICERs for comparison of platinum doublet chemotherapy versus pembrolizumab as first-line therapy for advanced NSCLC, under high level of study-to-study heterogeneity

	Chemotherapy		Pembrolizumab				
				Without End-of-Life adjustment		With End-of-Life adjustment	
	Cost ('000) USD (95% CrI)	QALY (95% CrI)	Cost ('000) USD (95% CrI)	QALY (95% CrI)	ICER ('000) USD per QALY (95% CrI)	QALY (95% CrI)	ICER ('000) USD per QALY (95% CrI)
No Dependency							
UK	34 (32-35)	1.03 (0.87-1.09)	99 (93-104)	1.89 (1.74-1.95)	77 (64-92)	3.01 (2.75-3.13)	33 (29-38)
US	70 (65-74)		132 (124-138)		73 (60-87)		31 (27-37)

C.2 High dependency scenario

In the third scenario, we incorporated a moderate dependency between each simulated patient's outcomes in the pembrolizumab and chemotherapy arms and between their associated OS and progression times by introducing a correlation of 0.9 between the arms and a correlation of 0.9 between progression and OS times via a Gaussian copula. In the chemotherapy arm, posterior mean QALYs gained remained 1.06 as in the moderate dependency scenario. The mean cost in the chemotherapy arm remained almost unchanged in both the UK and US setting. Mean cost in the pembrolizumab arm increased from

$121,000 to $123,000 in the UK setting, and from $160,000 to $164,000 in the US setting. In the absence of EoL adjustment, mean QALYs gained by patients on pembrolizumab decreased from 1.8 to 1.78, leading to ICER per QALY gained of $121,000 for the UK and $116,000 for the US. With EoL adjustment, ICERs per QALY gained for the UK and US setting were $62,000 and $60,000 respectively. See Table 22Table 22. The probability that pembrolizumab was cost-effective was 0.1% with respect to the UK (USD 42,048) threshold and 98.6% with respect to the US (USD 100,000) threshold in the presence of EoL adjustment. The probabilities were <1% and 9.3% with respect to the UK and US thresholds respectively in the absence of EoL adjustment. The results from this strong dependence scenario were similar to the moderate dependence scenario.

Table 22: Posterior mean (95% CrI) costs (2018 USD), QALYs and ICERs for comparison of platinum doublet chemotherapy versus pembrolizumab as first-line therapy for advanced NSCLC, under strong dependency between patients' hypothetical outcomes

	Chemotherapy		**Pembrolizumab**				
				Without End-of-Life adjustment		**With End-of-Life adjustment**	
	Cost ('000) USD (95% CrI)	**QALY (95% CrI)**	**Cost ('000) USD (95% CrI)**	**QALY (95% CrI)**	**ICER ('000) USD per QALY (95% CrI)**	**QALY (95% CrI)**	**ICER ('000) USD per QALY (95% CrI)**
High Dependency							
UK	38 (36-40)	1.06 (0.94-1.13)	123 (114-131)	1.78 (1.55-1.88)	121 (96-176)	2.45 (2.03-2.67)	62 (50-89)
US	82 (77-87)		164 (152-174)		116 (90-170)		60 (46-86)

C.3 Analysis incorporating discounting

We considered a discounting factor to adjust the costs and utilities for the pembrolizumab and chemotherapy arms to present values. We chose a discounting factor corresponding to 3% on an annual basis. Accumulated costs and utilities were multiplied by 1/(1+0.03)^n , where n is the corresponding year in which the cost or utility occurred. The results were qualitatively unchanged with slightly higher ICER values and lower probabilities of pembrolizumab being cost effective. The results can be found in Table 23.

Table 23: Posterior mean (95% CrI) costs (2018 USD), QALYs and ICERs for comparison of platinum doublet chemotherapy versus pembrolizumab as first-line therapy for advanced NSCLC

	Chemotherapy			Pembrolizumab			
				Without End-of-Life adjustment		With End-of-Life adjustment	
	Cost ('000) USD (95% CrI)	QALY (95% CrI)	Cost ('000) USD (95% CrI)	QALY (95% CrI)	ICER ('000) USD per QALY (95% CrI)	QALY (95% CrI)	ICER ('000) USD per QALY (95% CrI)
No Dependency							
UK	33 (31-35)	0.98 (0.88-1.05)	95 (86-104)	1.62 (1.43-1.69)	100 (82-140)	2.54 (2.18-2.68)	40 (34-52)
US	71 (67-76)		128 (116-139)		90 (71-127)		36 (29-47)
Moderate Dependency							
UK	37 (35-39)	0.95 (0.85-1.01)	116 (108-123)	1.53 (1.34-1.62)	138 (110-203)	2.25 (1.89-2.41)	61 (51-84)
US	79 (75-84)		154 (143-163)		130 (101-193)		58 (47-79)
High Dependency							
UK	37 (35-39)	0.94 (0.84-1.01)	118 (110-126)	1.52 (1.33-1.61)	143 (113-213)	2.13 (1.76-2.31)	70 (56-100)
US	80 (76-85)		157 (146-167)		136 (104-204)		66 (51-96)

C.4 Analysis based on parametric survival model (Weibull distribution)

We considered a parametric survival model using the Weibull distribution. We used the fitted parametric survival curves, to independently generate 10,000,000 progression and OS times and construct patient trajectories for the pembrolizumab and chemotherapy arms. The results are summarized in Table 24 below.

Table 24: Mean (95% CI) costs (2018 USD), QALYs and ICERs for comparison of platinum doublet chemotherapy versus pembrolizumab as first-line therapy for advanced NSCLC for Weibull survival model

	Chemotherapy			Pembrolizumab			
				Without End-of-Life adjustment		With End-of-Life adjustment	
	Cost ('000) USD (95% CrI)	QALY (95% CrI)	Cost ('000) USD (95% CrI)	QALY (95% CrI)	ICER ('000) USD per QALY (95% CrI)	QALY (95% CrI)	ICER ('000) USD per QALY (95% CrI)
No Dependency							
UK	36 (35-36)	1.15 (1.14 - 1.16)	103 (102-104)	1.97 (1.95-1.98)	82 (80-84)	3.11 (3.08-3.15)	34 (34-35)
US	77 (76-78)		137 (136-139)		74 (72-76)		31 (30-31)
Moderate Dependency							
UK	40 (39-40)	1.1 (1.1-1.11)	124 (123-125)	1.85 (1.83-1.86)	113 (110-116)	2.74 (2.71-2.77)	51 (50-52)
US	85 (85-86)		164 (163-165)		106 (102-109)		48 (47-49)
High Dependency							
UK	40 (40-41)	1.1 (1.09 -1.1)	126 (125-127)	1.83 (1.81-1.84)	117 (114-121)	2.51 (2.47-2.54)	61 (59-62)
US	86 (86-87)		167 (165-168)		110 (107-113)		57 (55-59)

REFERENCES

[1] K. Takahashi and S. Yamanaka, "Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors," *cell*, vol. 126, pp. 663-676, 2006.

[2] E. Hupp, S. Roy, and M. Watzke, "NASA Finds Direct Proof of Dark Matter ", ed. NASA: NASA, 2006.

[3] J. Markoff, "Google cars drive themselves, in traffic," *The New York Times*, vol. 10, p. 9, 2010.

[4] H. J. Burstein, L. Krilov, J. B. Aragon-Ching, N. N. Baxter, E. G. Chiorean, W. A. Chow, *et al.*, "Clinical cancer advances 2017: annual report on progress against cancer from the American Society of Clinical Oncology," *Journal of Clinical Oncology*, vol. 35, pp. 1341-1367, 2017.

[5] J. Heymach, L. Krilov, A. Alberg, N. Baxter, S. M. Chang, R. Corcoran, *et al.*, "Clinical Cancer Advances 2018: Annual Report on Progress Against Cancer From the American Society of Clinical Oncology," *Journal of Clinical Oncology*, vol. 36, pp. 1020-1044, 2018.

[6] Q. Ma, L. C. Dieterich, and M. Detmar, "Biology of Melanoma Metastasis," *Melanoma*, pp. 1-17, 2017.

[7] E. Bajetta, M. Del Vecchio, C. Bernard-Marty, M. Vitali, R. Buzzoni, O. Rixe, *et al.*, "Metastatic melanoma: Chemotherapy," *Seminars in Oncology*, vol. 29, pp. 427-445.

[8] A. Barth, L. A. Wanek, and D. L. Morton, "Prognostic factors in 1,521 melanoma patients with distant metastases," *Journal of the American College of Surgeons*, vol. 181, pp. 193-201, 1995.

[9] A. M. M. Eggermont, A. Spatz, and C. Robert, "Cutaneous melanoma," *The Lancet*, vol. 383, pp. 816-827, 2014/03/01/ 2014.

[10] T. K. Eigentler, U. M. Caroli, P. Radny, and C. Garbe, "Palliative therapy of disseminated malignant melanoma: a systematic review of 41 randomised clinical trials," *The Lancet Oncology*, vol. 4, pp. 748-759, 2003/12/01/ 2003.

[11] M. Harries, J. Malvehy, C. Lebbe, L. Heron, J. Amelio, Z. Szabo, *et al.*, "Treatment patterns of advanced malignant melanoma (stage III–IV) – A review of current standards in Europe," *European Journal of Cancer*, vol. 60, pp. 179-189, 2016/06/01/ 2016.

[12] M. R. Middleton, S. Dalle, J. Claveau, P. Mut, S. Hallmeyer, P. Plantin, *et al.*, "Real-world treatment practice in patients with advanced melanoma in the era before ipilimumab: results from the IMAGE study," *Cancer Medicine,* vol. 5, pp. 1436-1443, 2016.

[13] D. Schadendorf, F. S. Hodi, C. Robert, J. S. Weber, K. Margolin, O. Hamid, *et al.*, "Pooled Analysis of Long-Term Survival Data From Phase II and Phase III Trials of Ipilimumab in Unresectable or Metastatic Melanoma," *Journal of Clinical Oncology,* vol. 33, pp. 1889-1894, 2015.

[14] N. D. James, M. R. Spears, N. W. Clarke, D. P. Dearnaley, J. S. De Bono, J. Gale, *et al.*, "Survival with Newly Diagnosed Metastatic Prostate Cancer in the "Docetaxel Era": Data from 917 Patients in the Control Arm of the STAMPEDE Trial (MRC PR08, CRUK/06/019)," *European Urology,* vol. 67, pp. 1028-1038, 2015/06/01/ 2015.

[15] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2017," *CA: A Cancer Journal for Clinicians,* vol. 67, pp. 7-30, 2017.

[16] C. A. Heinlein and C. Chang, "Androgen Receptor in Prostate Cancer," *Endocrine Reviews,* vol. 25, pp. 276-308, 2004.

[17] H. I. Scher, K. Fizazi, F. Saad, M. E. Taplin, C. N. Sternberg, K. Miller, *et al.*, "Increased Survival with Enzalutamide in Prostate Cancer after Chemotherapy," *New England Journal of Medicine,* vol. 367, pp. 1187-1197, Sep 2012.

[18] J. S. De Bono, C. J. Logothetis, A. Molina, K. Fizazi, S. North, L. Chu, *et al.*, "Abiraterone and Increased Survival in Metastatic Prostate Cancer," *New England Journal of Medicine,* vol. 364, pp. 1995-2005, May 2011.

[19] D. E. Gerber and J. H. Schiller, "Maintenance Chemotherapy for Advanced Non–Small-Cell Lung Cancer: New Life for an Old Idea," *Journal of Clinical Oncology,* vol. 31, pp. 1009-1020, 2013.

[20] M. Reck, D. Rodríguez-Abreu, A. G. Robinson, R. Hui, T. Csőszi, A. Fülöp, *et al.*, "Pembrolizumab versus Chemotherapy for PD-L1–Positive Non–Small-Cell Lung Cancer," *New England Journal of Medicine,* vol. 375, pp. 1823-1833, 2016.

[21] R. S. Herbst, P. Baas, D.-W. Kim, E. Felip, J. L. Pérez-Gracia, J.-Y. Han, *et al.*, "Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial," *The Lancet,* vol. 387, pp. 1540-1550, 2016.

[22] A. A. Alsheikh-Ali, W. Qureshi, M. H. Al-Mallah, and J. P. Ioannidis, "Public availability of published research data in high-impact journals," *PloS one,* vol. 6, p. e24357, 2011.

[23] N. Draper and I. Guttman, "Bayesian Estimation of the Binomial Parameter," *Technometrics,* vol. 13, pp. 667-673, 1971/08/01 1971.

[24] I. Olkin, A. J. Petkau, and J. V. Zidek, "A Comparison of n Estimators for the Binomial Distribution," *Journal of the American Statistical Association,* vol. 76, pp. 637-642, 1981.

[25] S. Basu, "Ch. 31. Bayesian inference for the number of undetected errors," *Handbook of Statistics,* vol. 22, pp. 1131-1150, 2003.

[26] P. B. Chapman , A. Hauschild , C. Robert , J. B. Haanen , P. Ascierto , J. Larkin , *et al.*, "Improved Survival with Vemurafenib in Melanoma with BRAF V600E Mutation," *New England Journal of Medicine,* vol. 364, pp. 2507-2516, 2011.

[27] K. T. Flaherty , J. R. Infante , A. Daud , R. Gonzalez , R. F. Kefford , J. Sosman , *et al.*, "Combined BRAF and MEK Inhibition in Melanoma with BRAF V600 Mutations," *New England Journal of Medicine,* vol. 367, pp. 1694-1703, 2012.

[28] A. Hauschild, J.-J. Grob, L. V. Demidov, T. Jouary, R. Gutzmer, M. Millward, *et al.*, "Dabrafenib in BRAF-mutated metastatic melanoma: a multicentre, open-label, phase 3 randomised controlled trial," *The Lancet,* vol. 380, pp. 358-365, 2012/07/28/ 2012.

[29] G. V. Long , D. Stroyakovskiy , H. Gogas , E. Levchenko , F. de Braud , J. Larkin , *et al.*, "Combined BRAF and MEK Inhibition versus BRAF Inhibition Alone in Melanoma," *New England Journal of Medicine,* vol. 371, pp. 1877-1888, 2014.

[30] G. V. Long, D. Stroyakovskiy, H. Gogas, E. Levchenko, F. de Braud, J. Larkin, *et al.*, "Dabrafenib and trametinib versus dabrafenib and placebo for Val600 BRAF-mutant melanoma: a multicentre, double-blind, phase 3 randomised controlled trial," *The Lancet,* vol. 386, pp. 444-451, 2015/08/01/ 2015.

[31] S. Schubbert, K. Shannon, and G. Bollag, "Hyperactive Ras in developmental disorders and cancer," *Nature Reviews Cancer,* vol. 7, p. 295, 04/01/online 2007.

[32] S. Seton-Rogers, "Therapeutics: delving deeper into resistance," *Nat Rev Cancer,* vol. 14, p. 7, Jan 2014.

[33] J. S. Weber, S. P. D'Angelo, D. Minor, F. S. Hodi, R. Gutzmer, B. Neyns, *et al.*, "Nivolumab versus chemotherapy in patients with advanced melanoma who progressed after anti-CTLA-4 treatment (CheckMate 037): a randomised, controlled, open-label, phase 3 trial," *The Lancet Oncology,* vol. 16, pp. 375-384, 2015/04/01/ 2015.

[34] M. A. Postow, J. Chesney, A. C. Pavlick, C. Robert, K. Grossmann, D. McDermott, *et al.*, "Nivolumab and Ipilimumab versus Ipilimumab in Untreated Melanoma," *New England Journal of Medicine,* vol. 372, pp. 2006-2017, 2015.

[35] C. Robert , G. V. Long , B. Brady , C. Dutriaux , M. Maio , L. Mortier , *et al.*, "Nivolumab in Previously Untreated Melanoma without BRAF Mutation," *New England Journal of Medicine,* vol. 372, pp. 320-330, 2015.

[36] C. Robert , L. Thomas , I. Bondarenko , S. O'Day , J. Weber , C. Garbe , *et al.*, "Ipilimumab plus Dacarbazine for Previously Untreated Metastatic Melanoma," *New England Journal of Medicine,* vol. 364, pp. 2517-2526, 2011.

[37] R. H. I. Andtbacka, H. L. Kaufman, F. Collichio, T. Amatruda, N. Senzer, J. Chesney, *et al.*, "Talimogene Laherparepvec Improves Durable Response Rate in Patients With Advanced Melanoma," *Journal of Clinical Oncology,* vol. 33, pp. 2780-2788, 2015.

[38] D. Moher, A. Liberati, J. Tetzlaff, and D. G. Altman, "Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement," *International journal of surgery,* vol. 8, pp. 336-341, 2010.

[39] J. D. Wolchok, A. Hoos, S. O'Day, J. S. Weber, O. Hamid, C. Lebbé, *et al.*, "Guidelines for the Evaluation of Immune Therapy Activity in Solid Tumors: Immune-Related Response Criteria," *Clinical Cancer Research,* vol. 15, pp. 7412-7420, 2009.

[40] H. C. Bucher, G. H. Guyatt, L. E. Griffith, and S. D. Walter, "The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials," *Journal of clinical epidemiology,* vol. 50, pp. 683-691, 1997.

[41] V. Hasselblad, "Meta-analysis of Multitreatment Studies," *Medical Decision Making,* vol. 18, pp. 37-43, 1998.

[42] T. Lumley, "Network meta-analysis for indirect treatment comparisons," *Statistics in Medicine,* vol. 21, pp. 2313-2324, 2002.

[43] T. C. Smith, D. J. Spiegelhalter, and A. Thomas, "Bayesian approaches to random-effects meta-analysis: a comparative study," *Statistics in medicine,* vol. 14, pp. 2685-2699, 1995.

[44] G. Lu and A. E. Ades, "Combination of direct and indirect evidence in mixed treatment comparisons," *Statistics in Medicine,* vol. 23, pp. 3105-3124, 2004.

[45] R Core Team, "R: A Language and Environment for Statistical Computing," ed: R Foundation for Statistical Computing, 2016.

[46] K. Hornik, F. Leisch, and A. Zeileis, "JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling," in *Proceedings of DSC*, 2003, p. 1.1.

[47] M. Plummer, "rjags: Bayesian Graphical Models using MCMC," *URL:* http://cran.r-project.org/web/packages/rjags, vol. R package version 4-6, 2016.

[48] G. Salanti, A. E. Ades, and J. P. A. Ioannidis, "Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial," *Journal of Clinical Epidemiology,* vol. 64, pp. 163-171, 2011/02/01/ 2011.

[49] R. DerSimonian and N. Laird, "Meta-analysis in clinical trials," *Controlled clinical trials,* vol. 7, pp. 177-188, 1986.

[50] J. P. T. Higgins, D. G. Altman, P. C. Gøtzsche, P. Jüni, D. Moher, A. D. Oxman, *et al.*, "The Cochrane Collaboration's tool for assessing risk of bias in randomised trials," *BMJ,* vol. 343, 2011.

[51] F. Hodi, S. Lee, D. F. McDermott, and et al., "Ipilimumab plus sargramostim vs ipilimumab alone for treatment of metastatic melanoma: A randomized clinical trial," *JAMA,* vol. 312, pp. 1744-1753, 2014.

[52] A. Daud, J. S. Weber, J. A. Sosman, K. Kim, R. Gonzalez, O. Hamid, *et al.*, "Updated overall survival (OS) results for BRF113220, a phase I–II study of dabrafenib alone versus combined dabrafenib and trametinib in patients with BRAF V600 metastatic melanoma (MM)," *Journal of Clinical Oncology,* vol. 33, pp. 9036-9036, 2015.

[53] K. T. Flaherty , C. Robert , P. Hersey , P. Nathan , C. Garbe , M. Milhem , *et al.*, "Improved Survival with MEK Inhibition in BRAF-Mutated Melanoma," *New England Journal of Medicine,* vol. 367, pp. 107-114, 2012.

[54] A. Hauschild, J. Grobb, L. Demidov, T. Jouary, R. Gutzmer, M. Millward, *et al.*, "1092PDAN UPDATE ON OVERALL SURVIVAL (OS) AND FOLLOW-ON THERAPIES IN BREAK-3, A PHASE III, RANDOMIZED TRIAL: DABRAFENIB (D) VS. DACARBAZINE (DTIC) IN PATIENTS (PTS) WITH BRAF V600E MUTATION-POSITIVE METASTATIC MELANOMA (MM)," *Annals of Oncology,* vol. 25, pp. iv378-iv378, 2014.

[55] J. M. Kirkwood, L. Bastholt, C. Robert, J. Sosman, J. Larkin, P. Hersey, *et al.*, "Phase II, Open-Label, Randomized Trial of the MEK1/2 Inhibitor Selumetinib as Monotherapy versus Temozolomide in Patients with Advanced Melanoma," *Clinical Cancer Research,* vol. 18, pp. 555-567, 2012.

[56] J. Larkin , P. A. Ascierto , B. Dréno , V. Atkinson , G. Liszkay , M. Maio , *et al.*, "Combined Vemurafenib and Cobimetinib in BRAF-Mutated Melanoma," *New England Journal of Medicine,* vol. 371, pp. 1867-1876, 2014.

[57] J. Larkin, V. Chiarion-Sileni, R. Gonzalez, J. J. Grob, C. L. Cowey, C. D. Lao, *et al.*, "Combined Nivolumab and Ipilimumab or Monotherapy in Untreated Melanoma," *New England Journal of Medicine,* vol. 373, pp. 23-34, 2015.

[58] J. M. G. Larkin, Y. Yan, G. A. McArthur, P. A. Ascierto, G. Liszkay, M. Maio, *et al.*, "Update of progression-free survival (PFS) and correlative biomarker analysis

from coBRIM: Phase III study of cobimetinib (cobi) plus vemurafenib (vem) in advanced BRAF-mutated melanoma," *Journal of Clinical Oncology*, vol. 33, pp. 9006-9006, 2015.

[59] N. R. Latimer, H. Bell, K. R. Abrams, M. M. Amonkar, and M. Casey, "Adjusting for treatment switching in the METRIC study shows further improved overall survival with trametinib compared with chemotherapy," *Cancer Medicine*, vol. 5, pp. 806-815, 2016.

[60] G. A. McArthur, P. B. Chapman, C. Robert, J. Larkin, J. B. Haanen, R. Dummer, *et al.*, "Safety and efficacy of vemurafenib in BRAFV600E and BRAFV600K mutation-positive melanoma (BRIM-3): extended follow-up of a phase 3, randomised, open-label study," *The Lancet Oncology*, vol. 15, pp. 323-332, 2014/03/01/ 2014.

[61] A. Ribas, R. Kefford, M. A. Marshall, C. J. A. Punt, J. B. Haanen, M. Marmol, *et al.*, "Phase III Randomized Clinical Trial Comparing Tremelimumab With Standard-of-Care Chemotherapy in Patients With Advanced Melanoma," *Journal of Clinical Oncology*, vol. 31, pp. 616-622, 2013.

[62] A. Ribas, I. Puzanov, R. Dummer, D. Schadendorf, O. Hamid, C. Robert, *et al.*, "Pembrolizumab versus investigator-choice chemotherapy for ipilimumab-refractory melanoma (KEYNOTE-002): a randomised, controlled, phase 2 trial," *The Lancet Oncology*, vol. 16, pp. 908-918, 2015.

[63] C. Robert, R. Dummer, R. Gutzmer, P. Lorigan, K. B. Kim, M. Nyakas, *et al.*, "Selumetinib plus dacarbazine versus placebo plus dacarbazine as first-line treatment for BRAF-mutant metastatic melanoma: a phase 2 double-blind randomised study," *The Lancet Oncology*, vol. 14, pp. 733-740, 2013/07/01/ 2013.

[64] C. Robert, B. Karaszewska, J. Schachter, P. Rutkowski, A. Mackiewicz, D. Stroiakovski, *et al.*, "Improved Overall Survival in Melanoma with Combined Dabrafenib and Trametinib," *New England Journal of Medicine*, vol. 372, pp. 30-39, 2015.

[65] C. Robert, B. Karaszewska, J. Schachter, P. Rutkowski, A. Mackiewicz, D. Stroyakovskiy, *et al.*, "Two year estimate of overall survival in COMBI-v, a randomized, open-label, phase III study comparing the combination of dabrafenib (D) and trametinib (T) with vemurafenib (Vem) as first-line therapy in patients (pts) with unresectable or metastatic BRAF V600E/K mutation-positive cutaneous melanoma," *European Journal of Cancer*, vol. 51, p. S663, 2015.

[66] C. Robert, J. Schachter, G. V. Long, A. Arance, J. J. Grob, L. Mortier, *et al.*, "Pembrolizumab versus Ipilimumab in Advanced Melanoma," *New England Journal of Medicine*, vol. 372, pp. 2521-2532, 2015.

[67] M. Postow, J. Chesney, A. Pavlick, C. Robert, K. Grossmann, D. McDermott, *et al.*, "Abstract CT002: Initial report of overall survival rates from a randomized

phase II trial evaluating the combination of nivolumab (NIVO) and ipilimumab (IPI) in patients with advanced melanoma (MEL)," *Cancer Research,* vol. 76, pp. CT002-CT002, 2016.

[68] E. Simeone, G. Gentilcore, D. Giannarelli, A. M. Grimaldi, C. Caracò, M. Curvietto, *et al.*, "Immunological and biological changes during ipilimumab treatment and their potential correlation with clinical response and survival in patients with advanced melanoma," *Cancer Immunology, Immunotherapy,* vol. 63, pp. 675-683, July 01 2014.

[69] F. S. Hodi, W.-J. Hwu, R. Kefford, J. S. Weber, A. Daud, O. Hamid, *et al.*, "Evaluation of Immune-Related Response Criteria and RECIST v1.1 in Patients With Advanced Melanoma Treated With Pembrolizumab," *Journal of Clinical Oncology,* vol. 34, pp. 1510-1517, 2016.

[70] P. N. A. Jr, I. L. Santoro, H. Tadokoro, G. d. L. Lopes, B. A. Filardi, P. Oliveira, *et al.*, "The role of PD-L1 expression as a predictive biomarker in advanced non-small-cell lung cancer: a network meta-analysis," *Immunotherapy,* vol. 8, pp. 479-488, 2016.

[71] J. Larkin, C. D. Lao, W. J. Urba, and et al., "Efficacy and safety of nivolumab in patients with braf v600 mutant and braf wild-type advanced melanoma: A pooled analysis of 4 clinical trials," *JAMA Oncology,* vol. 1, pp. 433-440, 2015.

[72] M. J. Ratcliffe, A. Sharpe, A. Midha, C. Barker, P. Scorer, and J. Walker, "Abstract LB-094: A comparative study of PD-L1 diagnostic assays and the classification of patients as PD-L1 positive and PD-L1 negative," *Cancer Research,* vol. 76, pp. LB-094-LB-094, 2016.

[73] A. H. Scheel, M. Dietel, L. C. Heukamp, K. Jöhrens, T. Kirchner, S. Reu, *et al.*, "Diagnostic PD-L1 immunohistochemistry in NSCLC: Results of the first German harmonization study," *Journal of Clinical Oncology,* vol. 34, pp. 3028-3028, 2016.

[74] S. Averbuch, K. Emancipator, I. McCaffery, A. McElhinny, D. Stanforth, J. Walker, *et al.*, "A Blueprint Proposal for Companion Diagnostic Comparability," in *FDA-AACR-ASCO Public Workshop,* Washington, DC, 2015.

[75] T. Powles, J. P. Eder, G. D. Fine, F. S. Braiteh, Y. Loriot, C. Cruz, *et al.*, "MPDL3280A (anti-PD-L1) treatment leads to clinical activity in metastatic bladder cancer," *Nature,* vol. 515, p. 558, 11/26/online 2014.

[76] P. C. Tumeh, C. L. Harview, J. H. Yearley, I. P. Shintaku, E. J. M. Taylor, L. Robert, *et al.*, "PD-1 blockade induces responses by inhibiting adaptive immune resistance," *Nature,* vol. 515, p. 568, 11/26/online 2014.

[77] A. Ribas, M. Butler, J. Lutzky, D. P. Lawrence, C. Robert, W. Miller, *et al.*, "Phase I study combining anti-PD-L1 (MEDI4736) with BRAF (dabrafenib) and/or MEK

(trametinib) inhibitors in advanced melanoma," *Journal of Clinical Oncology*, vol. 33, pp. 3003-3003, 2015.

[78] Astellas Pharma Inc. (2014). *A Multi-center, Single Arm Study of Enzalutamide in Patients With Progressive Metastatic Castration-Resistant Prostate Cancer Previously Treated With Abiraterone Acetate*. Available: https://clinicaltrials.gov/ct2/show/study/NCT02116582

[79] S. Diem, B. Kasenda, J. Martin-Liberal, A. Lee, D. Chauhan, M. Gore, *et al.*, "Prognostic score for patients with advanced melanoma treated with ipilimumab," *European Journal of Cancer*, vol. 51, pp. 2785-2791, 2015/12/01/ 2015.

[80] S. Diem, B. Kasenda, L. Spain, J. Martin-Liberal, R. Marconcini, M. Gore, *et al.*, "Serum lactate dehydrogenase as an early marker for outcome in patients treated with anti-PD-1 therapy in metastatic melanoma," *British Journal Of Cancer*, vol. 114, p. 256, 01/21/online 2016.

[81] R. Dummer, D. Schadendorf, P. A. Ascierto, A. M. A. Fernández, C. Dutriaux, M. Maio, *et al.*, "Results of NEMO: A phase III trial of binimetinib (BINI) vs dacarbazine (DTIC) in NRAS-mutant cutaneous melanoma," in *2016 ASCO Annual Meeting*, 2016.

[82] A. Jemal, R. Siegel, J. Xu, and E. Ward, "Cancer Statistics, 2010," *CA: A Cancer Journal for Clinicians*, vol. 60, pp. 277-300, 2010.

[83] M. Borre, B. Nerstrøm, and J. Overgaard, "The Dilemma of Prostate Cancer: A Growing Human and Economic Burden Irrespective of Treatment Strategies," *Acta Oncologica*, vol. 36, pp. 681-687, 1997/01/01 1997.

[84] D. A. Loblaw, K. S. Virgo, R. Nam, M. R. Somerfield, E. Ben-Josef, D. S. Mendelson, *et al.*, "Initial Hormonal Management of Androgen-Sensitive Metastatic, Recurrent, or Progressive Prostate Cancer: 2007 Update of an American Society of Clinical Oncology Practice Guideline," *Journal of Clinical Oncology*, vol. 25, pp. 1596-1605, 2007.

[85] D. G. McLeod, "Hormonal therapy: historical perspective to future directions," *Urology*, vol. 61, pp. 3-7, 2003.

[86] S. Gillessen, A. Omlin, G. Attard, J. S. de Bono, E. Efstathiou, K. Fizazi, *et al.*, "Management of patients with advanced prostate cancer: recommendations of the St Gallen Advanced Prostate Cancer Consensus Conference (APCCC) 2015," *Annals of Oncology*, vol. 26, pp. 1589-1604, 2015.

[87] A. Heidenreich, P. J. Bastian, J. Bellmunt, M. Bolla, S. Joniau, T. van der Kwast, *et al.*, "EAU Guidelines on Prostate Cancer. Part II: Treatment of Advanced, Relapsing, and Castration-Resistant Prostate Cancer," *European Urology*, vol. 65, pp. 467-479, 2014.

[88] D. R. Berthold, G. R. Pond, F. Soban, R. d. Wit, M. Eisenberger, and I. F. Tannock, "Docetaxel Plus Prednisone or Mitoxantrone Plus Prednisone for Advanced Prostate Cancer: Updated Survival in the TAX 327 Study," *Journal of Clinical Oncology,* vol. 26, pp. 242-245, 2008.

[89] J. S. de Bono, S. Oudard, M. Ozguroglu, S. Hansen, J.-P. Machiels, I. Kocak, *et al.*, "Prednisone plus cabazitaxel or mitoxantrone for metastatic castration-resistant prostate cancer progressing after docetaxel treatment: a randomised open-label trial," *The Lancet,* vol. 376, pp. 1147-1154, 2010.

[90] P. W. Kantoff , C. S. Higano , N. D. Shore , E. R. Berger , E. J. Small , D. F. Penson , *et al.*, "Sipuleucel-T Immunotherapy for Castration-Resistant Prostate Cancer," *New England Journal of Medicine,* vol. 363, pp. 411-422, 2010.

[91] K. Fizazi, H. I. Scher, A. Molina, C. J. Logothetis, K. N. Chi, R. J. Jones, *et al.*, "Abiraterone acetate for treatment of metastatic castration-resistant prostate cancer: final overall survival analysis of the COU-AA-301 randomised, double-blind, placebo-controlled phase 3 study," *The Lancet Oncology,* vol. 13, pp. 983-992, 2012.

[92] C. J. Ryan, M. R. Smith, K. Fizazi, F. Saad, P. F. A. Mulders, C. N. Sternberg, *et al.*, "Abiraterone acetate plus prednisone versus placebo plus prednisone in chemotherapy-naive men with metastatic castration-resistant prostate cancer (COU-AA-302): final overall survival analysis of a randomised, double-blind, placebo-controlled phase 3 study," *The Lancet Oncology,* vol. 16, pp. 152-160, 2015.

[93] H. I. Scher , K. Fizazi , F. Saad , M.-E. Taplin , C. N. Sternberg , K. Miller , *et al.*, "Increased Survival with Enzalutamide in Prostate Cancer after Chemotherapy," *New England Journal of Medicine,* vol. 367, pp. 1187-1197, 2012.

[94] T. M. Beer , A. J. Armstrong , D. E. Rathkopf , Y. Loriot , C. N. Sternberg , C. S. Higano , *et al.*, "Enzalutamide in Metastatic Prostate Cancer before Chemotherapy," *New England Journal of Medicine,* vol. 371, pp. 424-433, 2014.

[95] C. Parker, S. Nilsson, D. Heinrich, S. I. Helle, J. M. O'Sullivan, S. D. Fosså, *et al.*, "Alpha Emitter Radium-223 and Survival in Metastatic Prostate Cancer," *New England Journal of Medicine,* vol. 369, pp. 213-223, 2013.

[96] Y. Chen, N. J. Clegg, and H. I. Scher, "Anti-androgens and androgen-depleting therapies in prostate cancer: new agents for an established target," *The Lancet Oncology,* vol. 10, pp. 981-991, 2009.

[97] Y. Loriot, D. Bianchini, E. Ileana, S. Sandhu, A. Patrikidou, C. Pezaro, *et al.*, "Antitumour activity of abiraterone acetate against metastatic castration-resistant prostate cancer progressing after docetaxel and enzalutamide (MDV3100)," *Annals of Oncology,* vol. 24, pp. 1807-1812, 2013.

[98] K. L. Noonan, S. North, R. L. Bitting, A. J. Armstrong, S. L. Ellard, and K. N. Chi, "Clinical activity of abiraterone acetate in patients with metastatic castration-resistant prostate cancer progressing after enzalutamide," *Annals of Oncology,* vol. 24, pp. 1802-1807, 2013.

[99] S. Badrising, V. van der Noort, I. M. van Oort, H. P. van den Berg, M. Los, P. Hamberg, *et al.*, "Clinical activity and tolerability of enzalutamide (MDV3100) in patients with metastatic, castration-resistant prostate cancer who progress after docetaxel and abiraterone treatment," *Cancer,* vol. 120, pp. 968-975, 2014.

[100] D. Bianchini, D. Lorente, A. Rodriguez-Vida, A. Omlin, C. Pezaro, R. Ferraldeschi, *et al.*, "Antitumour activity of enzalutamide (MDV3100) in patients with metastatic castration-resistant prostate cancer (CRPC) pre-treated with docetaxel and abiraterone," *European Journal of Cancer,* vol. 50, pp. 78-84, 2013.

[101] K. Brasso, F. B. Thomsen, A. J. Schrader, S. C. Schmid, D. Lorente, M. Retz, *et al.*, "Enzalutamide Antitumour Activity Against Metastatic Castration-resistant Prostate Cancer Previously Treated with Docetaxel and Abiraterone: A Multicentre Analysis," *European Urology,* vol. 68, pp. 317-324, 2015.

[102] F. Petrelli, A. Coinu, K. Borgonovo, M. Cabiddu, M. Ghilardi, V. Lonati, *et al.*, "Enzalutamide After Docetaxel and Abiraterone Acetate Treatment in Prostate Cancer: A Pooled Analysis of 10 Case Series," *Clinical Genitourinary Cancer,* vol. 13, pp. 193-198, 2015.

[103] S. C. Schmid, A. Geith, A. Böker, R. Tauber, A. K. Seitz, M. Kuczyk, *et al.*, "Enzalutamide After Docetaxel and Abiraterone Therapy in Metastatic Castration-Resistant Prostate Cancer," *Advances in Therapy,* vol. 31, pp. 234-241, February 01 2014.

[104] A. J. Schrader, M. Boegemann, C.-H. Ohlmann, T. J. Schnoeller, L.-M. Krabbe, T. Hajili, *et al.*, "Enzalutamide in Castration-resistant Prostate Cancer Patients Progressing After Docetaxel and Abiraterone," *European Urology,* vol. 65, pp. 30-36, 2014.

[105] A. A. Azad, B. J. Eigl, R. N. Murray, C. Kollmannsberger, and K. N. Chi, "Efficacy of Enzalutamide Following Abiraterone Acetate in Chemotherapy-naive Metastatic Castration-resistant Prostate Cancer Patients," *European Urology,* vol. 67, pp. 23-29, 2015.

[106] H. H. Cheng, R. Gulati, A. Azad, R. Nadal, P. Twardowski, U. N. Vaishampayan, *et al.*, "Activity of enzalutamide in men with metastatic castration-resistant prostate cancer is affected by prior treatment with abiraterone and/or docetaxel," *Prostate Cancer And Prostatic Disease,* vol. 18, p. 122, 01/20/online 2015.

[107] T. Zhang, M. S. Dhawan, P. Healy, D. J. George, M. R. Harrison, J. Oldan, *et al.*, "Exploring the Clinical Benefit of Docetaxel or Enzalutamide After Disease Progression During Abiraterone Acetate and Prednisone Treatment in Men With

Metastatic Castration-Resistant Prostate Cancer," *Clinical Genitourinary Cancer,* vol. 13, pp. 392-399, 2015.

[108] E. S. Antonarakis , C. Lu , H. Wang , B. Luber , M. Nakazawa , J. C. Roeser , *et al.*, "AR-V7 and Resistance to Enzalutamide and Abiraterone in Prostate Cancer," *New England Journal of Medicine,* vol. 371, pp. 1028-1038, 2014.

[109] A. Agresti and B. A. Coull, "Approximate is Better than “Exact” for Interval Estimation of Binomial Proportions," *The American Statistician,* vol. 52, pp. 119-126, 1998/05/01 1998.

[110] A. Rohatgi. (2017). *WebPlotDigitizer (4.0 ed.).* Available: https://automeris.io/WebPlotDigitizer

[111] G. Gravis, J.-M. Boher, F. Joly, M. Soulié, L. Albiges, F. Priou, *et al.*, "Androgen Deprivation Therapy (ADT) Plus Docetaxel Versus ADT Alone in Metastatic Non castrate Prostate Cancer: Impact of Metastatic Burden and Long-term Survival Analysis of the Randomized Phase 3 GETUG-AFU15 Trial," *European Urology,* vol. 70, pp. 256-262, 2015.

[112] C. J. Sweeney, Y.-H. Chen, M. Carducci, G. Liu, D. F. Jarrard, M. Eisenberger, *et al.*, "Chemohormonal Therapy in Metastatic Hormone-Sensitive Prostate Cancer," *New England Journal of Medicine,* vol. 373, pp. 737-746, 2015.

[113] N. D. James, M. R. Sydes, N. W. Clarke, M. D. Mason, D. P. Dearnaley, M. R. Spears, *et al.*, "Addition of docetaxel, zoledronic acid, or both to first-line long-term hormone therapy in prostate cancer (STAMPEDE): survival results from an adaptive, multiarm, multistage, platform randomised controlled trial," *The Lancet,* vol. 387, pp. 1163-1177, 2015.

[114] Alliance for Clinical Trials in Oncology. (2014). *Phase III Trial of Enzalutamide (NSC# 766085) Versus Enzalutamide, Abiraterone and Prednisone for Castration Resistant Metastatic Prostate Cancer*. Available: https://www.clinicaltrials.gov/ct2/show/NCT01949337

[115] British Columbia Cancer Agency. (2014). *A Randomized Phase II Study of Sequencing Abiraterone Acetate and Enzalutamide in Metastatic Castration-Resistant Prostate Cancer*. Available: https://www.ClinicalTrials.gov/ct2/show/NCT02125357

[116] E. Efstathiou, M. A. Titus, S. Wen, A. SanMiguel, A. Hoang, A. D. Haas-Amatsaleh, *et al.*, "Enzalutamide (ENZA) in combination with abiraterone acetate (AA) in bone metastatic castration resistant prostate cancer (mCRPC)," *Journal of Clinical Oncology,* vol. 32, pp. 5000-5000, 2014.

[117] V. T. Devita Jr, T. Lawrence, and S. A. Rosenberg, *Cancer Principles and Practice of Oncology,* 10 ed.: Lippincott Williams & Wilkins, 2014.

[118] L. V. Sequist, J. C.-H. Yang, N. Yamamoto, K. O'Byrne, V. Hirsh, T. Mok, *et al.*, "Phase III Study of Afatinib or Cisplatin Plus Pemetrexed in Patients With Metastatic Lung Adenocarcinoma With EGFR Mutations," *Journal of Clinical Oncology,* vol. 31, pp. 3327-3334, 2013.

[119] T. S. Mok , Y.-L. Wu , S. Thongprasert , C.-H. Yang , D.-T. Chu , N. Saijo , *et al.*, "Gefitinib or Carboplatin–Paclitaxel in Pulmonary Adenocarcinoma," *New England Journal of Medicine,* vol. 361, pp. 947-957, 2009.

[120] R. Rosell, E. Carcereny, R. Gervais, A. Vergnenegre, B. Massuti, E. Felip, *et al.*, "Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non-small-cell lung cancer (EURTAC): a multicentre, open-label, randomised phase 3 trial," *The Lancet Oncology,* vol. 13, pp. 239-246, 2012.

[121] B. J. Solomon, T. Mok, D.-W. Kim, Y.-L. Wu, K. Nakagawa, T. Mekhail, *et al.*, "First-Line Crizotinib versus Chemotherapy in ALK-Positive Lung Cancer," *New England Journal of Medicine,* vol. 371, pp. 2167-2177, 2014.

[122] D. M. Parkin, F. Bray, J. Ferlay, and P. Pisani, "Global Cancer Statistics, 2002," *CA: A Cancer Journal for Clinicians,* vol. 55, pp. 74-108, 2005.

[123] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2016," *CA: A Cancer Journal for Clinicians,* vol. 66, pp. 7-30, 2016.

[124] J. Ferlay, I. Soerjomataram, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, *et al.*, "Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012," *International Journal of Cancer,* vol. 136, pp. E359-E386, 2015.

[125] J. D. Patel, M. A. Socinski, E. B. Garon, C. H. Reynolds, D. R. Spigel, M. R. Olsen, *et al.*, "PointBreak: A Randomized Phase III Study of Pemetrexed Plus Carboplatin and Bevacizumab Followed by Maintenance Pemetrexed and Bevacizumab Versus Paclitaxel Plus Carboplatin and Bevacizumab Followed by Maintenance Bevacizumab in Patients With Stage IIIB or IV Nonsquamous Non–Small-Cell Lung Cancer," *Journal of Clinical Oncology,* vol. 31, pp. 4349-4357, 2013.

[126] F. Barlesi, A. Scherpereel, V. Gorbunova, R. Gervais, A. Vikström, C. Chouaid, *et al.*, "Maintenance bevacizumab–pemetrexed after first-line cisplatin–pemetrexed–bevacizumab for advanced nonsquamous nonsmall-cell lung cancer: updated survival analysis of the AVAPERL (MO22089) randomized phase III trial," *Annals of Oncology,* vol. 25, pp. 1044-1052, 2014.

[127] A. Snyder, V. Makarov, T. Merghoub, J. Yuan, J. M. Zaretsky, A. Desrichard, *et al.*, "Genetic Basis for Clinical Response to CTLA-4 Blockade in Melanoma," *New England Journal of Medicine,* vol. 371, pp. 2189-2199, 2014.

[128] N. McGranahan, A. J. S. Furness, R. Rosenthal, S. Ramskov, R. Lyngaa, S. K. Saini, *et al.*, "Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade," *Science,* vol. 351, pp. 1463-1469, 2016.

[129] J. Brahmer, K. L. Reckamp, P. Baas, L. Crinò, W. E. E. Eberhardt, E. Poddubskaya, *et al.*, "Nivolumab versus Docetaxel in Advanced Squamous-Cell Non–Small-Cell Lung Cancer," *New England Journal of Medicine,* vol. 373, pp. 123-135, 2015.

[130] H. Borghaei, L. Paz-Ares, L. Horn, D. R. Spigel, M. Steins, N. E. Ready, *et al.*, "Nivolumab versus Docetaxel in Advanced Nonsquamous Non–Small-Cell Lung Cancer," *New England Journal of Medicine,* vol. 373, pp. 1627-1639, 2015.

[131] A. Rittmeyer, F. Barlesi, D. Waterkamp, K. Park, F. Ciardiello, J. von Pawel, *et al.*, "Atezolizumab versus docetaxel in patients with previously treated non-small-cell lung cancer (OAK): a phase 3, open-label, multicentre randomised controlled trial," *The Lancet,* vol. 389, pp. 255-265, 2017.

[132] A. Andrews, "Treating with Checkpoint Inhibitors—Figure $1 Million per Patient," *American Health & Drug Benefits,* vol. 8, pp. 9-9, 2015.

[133] J. Brahmer, et al, "OA 17.06 Updated Analysis of KEYNOTE-024: Pembrolizumab vs Platinum-Based Chemotherapy for Advanced NSCLC With PD-L1 TPS≥ 50%," *Journal of Thoracic Oncology,* vol. 12, pp. S1793-S1794, 2017.

[134] P. Guyot, A. E. Ades, M. J. N. M. Ouwens, and N. J. Welton, "Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves," *BMC Medical Research Methodology,* vol. 12, p. 9, 2012/02/01 2012.

[135] V. T. Farewell, "The use of mixture models for the analysis of survival data with long-term survivors," *Biometrics,* pp. 1041-1046, 1982.

[136] T. S. Ferguson, "A Bayesian Analysis of Some Nonparametric Problems," *Ann. Statist.,* vol. 1, pp. 209-230, 1973/03 1973.

[137] C. E. Antoniak, "Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems," *The annals of statistics,* pp. 1152-1174, 1974.

[138] V. Susarla and J. Van Ryzin, "Nonparametric Bayesian estimation of survival curves from incomplete observations," *Journal of the American Statistical Association,* vol. 71, pp. 897-902, 1976.

[139] J. Blum and V. Susarla, "On the posterior distribution of a Dirichlet process given randomly right censored observations," *Stochastic Processes and their Applications,* vol. 5, pp. 207-211, 1977.

[140] Y. W. Teh, M. I. Jordan, M. J. Beal, and D. M. Blei, "Hierarchical Dirichlet Processes," *Journal of the American Statistical Association,* vol. 101, pp. 1566-1581, 2006/12/01 2006.

[141] NICE, "Single Technology Appraisal Nivolumab for previously treated locally advanced or metastatic squamous non- small-cell lung cancer " *Committee Papers*, 2015.

[142] *GoodRx*. Available: www.goodrx.com

[143] NICE, "Ipilimumab for previously untreated advanced (unresectable or metastatic) melanoma," *Technology appraisal guidance*, 2014.

[144] NICE, "Pembrolizumab for untreated PD-L1-positive metastatic non-small-cell lung cancer," *Technology appraisal guidance*, 2017.

[145] W. Wong, Y. M. Yim, A. Kim, M. Cloutier, M. Gauthier-Loiselle, P. Gagnon-Sanschagrin, *et al.*, "Assessment of costs associated with adverse events in patients with cancer," *PLoS ONE,* vol. 13, 2018.

[146] S. Hurvitz, A. Guerin, M. Brammer, E. Guardino, Z.-Y. Zhou, D. L. Viau, *et al.*, "Investigation of adverse-event-related costs for patients with metastatic breast cancer in a real-world setting," *The oncologist,* vol. 19, pp. 901-908, 2014.

[147] B. Nafees, M. Stafford, S. Gavriel, S. Bhalla, and J. Watkins, "Health state utilities for non small cell lung cancer," *Health and Quality of Life Outcomes,* vol. 6, p. 84, October 21 2008.

[148] G. Stewart, L. Eddowes, L. Hamerslag, and J. Kusel, "The Impact Of Nice's End-Of-Life Threshold On Patient Access To New Cancer Therapies In England And Wales," *Value in Health,* vol. 17, p. A6, 2014.

[149] S. Dixon, L. Longworth, and A. Wailoo, "Assessing technologies at the end of life: a review of empirical evidence," 2009.

[150] NICE. (2017). *Nivolumab now available for lung cancer after company offers NICE new CDF deal*. Available: www.nice.org.uk/news/article/nivolumab-now-available-for-lung-cancer-after-company-offers-nice-new-cdf-deal

[151] M. Prada, M. Ruggeri, C. Sansone, D. De Fazio, A. Tettamanti, and M. Mantovani, "Timeline of authorization and reimbursement for oncology drugs in Italy in the last 3 years," *Medicine Access@ Point of Care,* vol. 1, p. maapoc. 0000007, 2017.

[152] R. Dickson, A. Boland, A. Bagust, M. Blundell, G. Massey, and Y. Dundar, "Ipilimumab for previously treated unresectable malignant melanoma: A Single Technology Appraisal," *Liverpool, UK: Liverpool Reviews and Implementation Group, The University of Liverpool*, 2011.

[153] G. Adunlin, V. Diaby, A. J. Montero, and H. Xiao, "Multicriteria decision analysis in oncology," *Health Expectations,* vol. 18, pp. 1812-1826, 2015.

[154] A. Davies, A. Briggs, J. Schneider, A. Levy, O. Ebeid, S. Wagner, *et al.*, "The ends justify the mean: outcome measures for estimating the value of new cancer therapies," *Health Outcomes Research in Medicine,* vol. 3, pp. e25-e36, 2012.

[155] R. Hettle, M. Corbett, S. Hinde, R. Hodgson, J. Jones-Diette, N. Woolacott, *et al.*, "The assessment and appraisal of regenerative medicines and cell therapy products: an exploration of methods for review, economic evaluation and appraisal," *Health technology assessment,* pp. 1-204, 2017.

[156] D. A. Goldstein, U. Bilal, and V. Prasad, "Pembrolizumab as first-line therapy in programmed death ligand 1–positive advanced lung cancer: Is it as effective as we think it is?," *Cancer,* vol. 123, pp. 3872-3874, 2017.

[157] M. Huang, Y. Lou, J. Pellissier, T. Burke, F. X. Liu, R. Xu, *et al.*, "Cost Effectiveness of Pembrolizumab vs. Standard-of-Care Chemotherapy as First-Line Treatment for Metastatic NSCLC that Expresses High Levels of PD-L1 in the United States," *PharmacoEconomics,* vol. 35, pp. 831-844, August 01 2017.

[158] L. Whitaker, "On the Poisson law of small numbers," *Biometrika,* vol. 10, pp. 36-71, 1914.

[159] R. A. Fisher, "The negative binomial distribution," *Annals of Eugenics,* vol. 11, pp. 182-187, 1941.

[160] J. B. S. Haldane, " The fitting of binomial distributions," *Annals of Eugenics,* vol. 11, pp. 179-181, 1941.

[161] R. J. Carroll and F. Lombard, "A Note on N Estimators for the Binomial Distribution," *Journal of the American Statistical Association,* vol. 80, pp. 423-426, 1985/06/01 1985.

[162] G. Casella, "Stabilizing Binomial n Estimators," *Journal of the American Statistical Association,* vol. 81, pp. 172-175, 1986/03/01 1986.

[163] S. Blumenthal and R. C. Dahiya, "Estimating the Binomial Parameter n," *Journal of the American Statistical Association,* vol. 76, pp. 903-909, 1981/12/01 1981.

[164] B. G. Lindsay, "Errors in inspection: integer parameter maximum likelihood in a finite population," *Journal of the American Statistical Association,* vol. 80, pp. 879-885, 1985.

[165] P. Hall, "On the erratic behavior of estimators of N in the binomial N, p distribution," *Journal of the American Statistical Association,* vol. 89, pp. 344-352, 1994.

[166] W. Kühne, P. Neumann, D. Stoyan, and H. Stoyan, "Pairs of successes in Bernoulli trials and a new n-estimator for the binomial distribution," *Applicationes Mathematicae,* vol. 22, pp. 331-337, 1994.

[167] A. DasGupta and H. Rubin, "Estimation of binomial parameters when both n, p are unknown," *Journal of Statistical Planning and Inference,* vol. 130, pp. 391-404, 2005.

[168] D. Feldman and M. Fox, "Estimation of the Parameter n in the Binomial Distribution," *Journal of the American Statistical Association,* vol. 63, pp. 150-158, 1968.

[169] A. E. Raftery, "Inference for the Binomial N parameter: A Bayes Empirical Bayes," 1986.

[170] W. D. Kahn, "A Cautionary Note for Bayesian Estimation of the Binomial Parameter n," *The American Statistician,* vol. 41, pp. 38-40, 1987/02/01 1987.

[171] G. Hamedani and G. Walter, "Bayes estimation of the binomial parameter n," *Communications in Statistics-Theory and Methods,* vol. 17, pp. 1829-1843, 1988.

[172] E. Günel and D. Chilko, "Estimation of parameter n of the binomial distribution," *Communications in Statistics-Simulation and Computation,* vol. 18, pp. 537-551, 1989.

[173] H. A. Bayoud, "Bayes and Empirical Bayes Estimation of the Parameter n in a Binomial Distribution," *Communications in Statistics-Simulation and Computation,* vol. 40, pp. 1422-1433, 2011.

[174] A. J. Hunter and H. J. Griffiths, "Bayesian Approach to Estimation of Insect Population Size," *Technometrics,* vol. 20, pp. 231-234, 1978/08/01 1978.

[175] S. Sadooghi-Alvandi, "Admissible estimation of the binomial parameter n," *The Annals of Statistics,* pp. 1634-1641, 1986.

[176] G. Zou and A. T. Wan, "Admissible and minimax estimation of the parameter n in the binomial distribution," *Journal of statistical planning and inference,* vol. 113, pp. 451-466, 2003.

[177] G. Iliopoulos, "Some new estimators of the binomial parameter n," *Communications in Statistics-Theory and Methods,* vol. 32, pp. 1361-1372, 2003.

[178] S. K. De and S. Zacks, "Two-stage and sequential estimation of parameter N of binomial distribution when p is known," *Sequential Analysis,* vol. 35, pp. 440-452, 2016.

[179] Student, "AN EXPLANATION OF DEVIATIONS FROM POISSON'S LAW IN PRACTICE," *Biometrika,* vol. 12, pp. 211-215, 1919.

[180] F. Binet, "The fitting of the positive binomial distribution when both parameters are estimated from the sample," *Annals of Human Genetics,* vol. 17, pp. 117-119, 1952.

[181] J. Skellam, "A probability distribution derived from the binomial distribution by regarding the probability of success as variable between the sets of trials," *Journal of the Royal Statistical Society. Series B (Methodological),* vol. 10, pp. 257-261, 1948.

[182] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, "A Limited Memory Algorithm for Bound Constrained Optimization," *SIAM Journal on Scientific Computing,* vol. 16, pp. 1190-1208, 1995.

[183] C. G. Broyden, "The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations," *IMA Journal of Applied Mathematics,* vol. 6, pp. 76-90, 1970.

[184] R. Fletcher, "A new approach to variable metric algorithms," *The Computer Journal,* vol. 13, pp. 317-322, 1970.

[185] D. Goldfarb, "A family of variable-metric methods derived by variational means," *Mathematics of computation,* vol. 24, pp. 23-26, 1970.

[186] D. F. Shanno, "Conditioning of quasi-Newton methods for function minimization," *Mathematics of computation,* vol. 24, pp. 647-656, 1970.

[187] M. Gilli, D. Maringer, and E. Schumann, *Numerical Methods and Optimization in Finance*: Academic Press, 2011.

[188] E. Schumann, "Numerical Methods and Optimization in Finance (NMOF) Manual. Package version 1.2-2)," 2011--2017.

[189] T. R. Tubman, M. D. Shields, B. G. Craig, H. C. Mulholland, and N. C. Nevin, "Congenital heart disease in Down's syndrome: two year prospective early screening study," *British Medical Journal,* vol. 302, pp. 1425-1427, 1991.