

A Network Meta Analysis of Treatments for Advanced/Metastatic Pancreatic Cancer

Matthew Connor Knowles



School of Health and Related Research,
The University of Sheffield,
United Kingdom

Supervisor: Dr. Kate Ren
September 2024
Word Count:
60 Credits

*A dissertation submitted in partial fulfillment of the requirements for the degree of Master of Science in
Statistics with Medical Applications*

Abstract

A modified Multilevel Network Meta-Regression (ML-NMR) was performed on data from eight studies comparing Gemcitabine (GEM) with six combination therapies and one standalone therapy. The comparative efficacy of the combination therapies was assessed in terms of the median overall survival (OS) and restricted mean survival time (RMST) estimates. GEM in combination with nab-paclitaxel (NAB) or capecitabine (CAP) provides improved median OS and RMST estimates for the treatment of advanced/metastatic pancreatic cancer compared to GEM monotherapy.

Acknowledgements

I would like to thank Dr Kate Ren for her invaluable supervision throughout this project. In addition, to my colleagues at OPEN Health who I have learnt so much from, and am continually inspired by. Finally, thank you to my friends and family for putting up with me talking about network meta analysis for the past two years.

For my late grandfather, Norman Fredrick Taylor (July 1943 - April 2015).



"But they that wait upon the Lord shall renew their strength; they shall mount up with wings as eagles; they shall run, and not be weary; and they shall walk, and not faint." Isaiah 40:31

CONTENTS

1	Background	6
1.1	Pancreatic Cancer	6
1.2	Treatment Landscape	6
1.2.1	Locally Advanced Pancreatic Cancer	7
1.2.2	Metastatic Pancreatic Cancer	8
1.3	Project Aim	8
1.4	Dissertation Structure	9
2	Survival Analysis Bakground	10
2.1	Survival Functions	10
2.2	The Kaplan-Meier Estimator and Reconcstructing Patient-Level Data	11
2.2.1	The Kaplam-Meier Estimator	11
2.3	Reconstructing Survival Data from Published Curves	12
2.4	Regression Models for Survival	12
2.4.1	Accelerated Failure Time Models	12
2.4.2	Proportional Hazards Models	12
2.5	Key Survival Metrics	13
2.5.1	The Hazard Ratio	13
2.5.2	Restriced Mean Survival Time	13
2.5.3	Median Survival	13
2.6	Parametric Models for Survival Analysis	14
2.6.1	Model Setup	14
2.6.2	Fitting Models	14
2.6.2.1	Right Censoring	14
2.6.2.2	Interval Censoring	14
3	Network Meta Analysis Theory	15
3.1	Building a Network of Evidence	15
3.2	Standard NMA Model	15
3.3	Multilevel Network Meta-Regression	16
3.4	Survival ML-NMR	17
3.5	Population-average estimates	19
3.5.1	Survival function	19
3.5.2	Hazard function	19
3.5.3	RMST	19
3.5.4	Median OS	19
3.6	Model Selection and Convergence	20
4	Included Studies	21
4.1	Study Eligibility Criteria	22
4.1.1	Eligibility by study	22
4.1.2	Analysis of eligibility	26
4.2	Covariates	26
4.3	Parametric Model Fitting	26

5 A Review of NMAs for Pancreatic Cancer	31
6 NMA of Pancreatic Cancer Trials	32
6.1 Network of Evidence	32
6.2 Model Fitting and Selection	32
6.3 Results	33
7 Conclusion and Discussions	35
7.1 Conclusion	35
7.2 Discussion and Further Work	35
7.3 Considerations for the ISPOR Good Practice Task Force	35
A Additional NMA Results	41
A.1 Model Convergence	41
B ISPOR Good Practice Questions	43
B.1 Evidence Base	43
B.2 Analysis	44
B.3 Reporting quality and transparency	44
B.4 Interpretation	44
B.5 Conflicts of Interest	44
C The PCNMA Package	46
C.1 Survival Functions	46
C.2 NMA Functions	46

LIST OF FIGURES

1.1	Pancreatic cancer incidence in the devolved nations 2017 to 2019	7
1.2	Pancreatic cancer incidence by sex, 1993 to 2018	8
3.1	Visualisation of combining trials into a network of evidence	16
4.1	Forest plot for median OS of GEM in each study	22
4.2	Forest plot for median OS of the comparator in each study	23
4.3	KM curves for each study	24
4.4	Log-cumulative hazard plot for each study	25
4.5	Cunningham (2009) parametric model extrapolations	27
4.6	Conroy (2011) parametric model extrapolations	27
4.7	Goldstein (2015) parametric model extrapolations	28
4.8	Goncalves (2012) parametric model extrapolations	28
4.9	Kindler (2011) parametric model extrapolations	29
4.10	Oettle (2005) parametric model extrapolations	29
4.11	Rocha Lima (2004) parametric model extrapolations	30
4.12	Spano (2008) parametric model extrapolations	30
6.1	Network of evidence	32
6.2	OS of each treatment in each population	33
6.3	Hazards of each treatment in each population	33
6.4	RMST of each treatment in each population	34
6.5	Median OS of each treatment in each population	34
A.1	Trace plot for the FE log-normal model ML-NMR	41
A.2	Trace plot for the FE log-normal model ML-NMR	42
C.1	Fitted Models as in Flexsurv	59
C.2	Fitted Models with <i>fit_distribution</i>	59

BACKGROUND

1.1 Pancreatic Cancer

Pancreatic cancer is the 10th most common cancer in the UK, accounting for around 3% of all new cancer cases [Cancer Research UK, 2024]. Pancreatic cancer has a particularly poor prognosis, with 9,558 deaths from 10,786 cases between 2017 and 2019. Figure 1.1 shows where the cases were across the UK in that time period. The relative incidence with respect to the overall number of cancer diagnoses is presented, and ranges between 2.53% in Scotland, to 2.84% in England. Only 3% of patients survive for more than five years [National Institute for Health and Care Excellence, 2018]. Part of the reason for the poor prognosis is that pancreatic cancer is hard to detect at early stages, meaning most people who present with symptoms already have advanced-stage pancreatic cancer by the time they present. Often, patients only notice symptoms when the tumour has spread to surrounding tissues, or metastasises to other organs [Kelsen et al., 1997]. The liver is the most common site of pancreatic cancer metastases [Deeb et al., 2015]. Common symptoms of pancreatic cancer include indigestion, stomach or back pain, loss of appetite and jaundice [Pancreatic Cancer UK, 2022]. These symptoms are common in other illnesses, which contributes to patients overlooking the fact their symptoms are consistent with pancreatic cancer.

Risk factors of pancreatic cancer include smoking, diabetes, obesity, and high-fat diets. Smoking is the dominant risk factor, with around 20% of cases being caused by cigarette smoking. In addition, cancers from smokers contain more genetic mutations when compared to cancers from non-smokers [Blackford et al., 2009].

Figure 1.2 presents the incidence of pancreatic cancer in the UK by sex from 1993 to 2018. SOURCE! The incidence rates were relatively flat until the early 2000s, at which point incidence began to increase. In the 2011-2018, the overall incidence (red line) plateaued and began to drop slightly, primarily driven by a declining incidence of pancreatic cancer in females from around 2013. The incidence has continued to rise in males in this time period. Some of this increase in incidence can be accounted for by advances in diagnostic technologies, such as Positron Emission Tomography (PET) scans and Endoscopic Ultrasound (EUS).

1.2 Treatment Landscape

Gemcitabine (GEM) ($C_9H_{11}F_2N_3O_4$) is a standard first-line treatment for pancreatic cancer administered intravenously [National Institute for Health and Care Excellence, 2018]. It was approved for the UK market in 1997. GEM is also used to treat other types of cancer, including breast cancer, bladder cancer and non-small-cell lung cancer [Wong et al., 2009]. GEM can be administered alone or in combination with another medication.

This dissertation considered six treatments that were given in combination with GEM: capecitabine (CAP) ($C_{15}H_{22}FN_3O_6$), axitinib (AXI) ($C_{22}H_{18}N_4OS$), pemetrexed (PEM) ($C_{20}H_{21}N_5O_6$), sorafenib

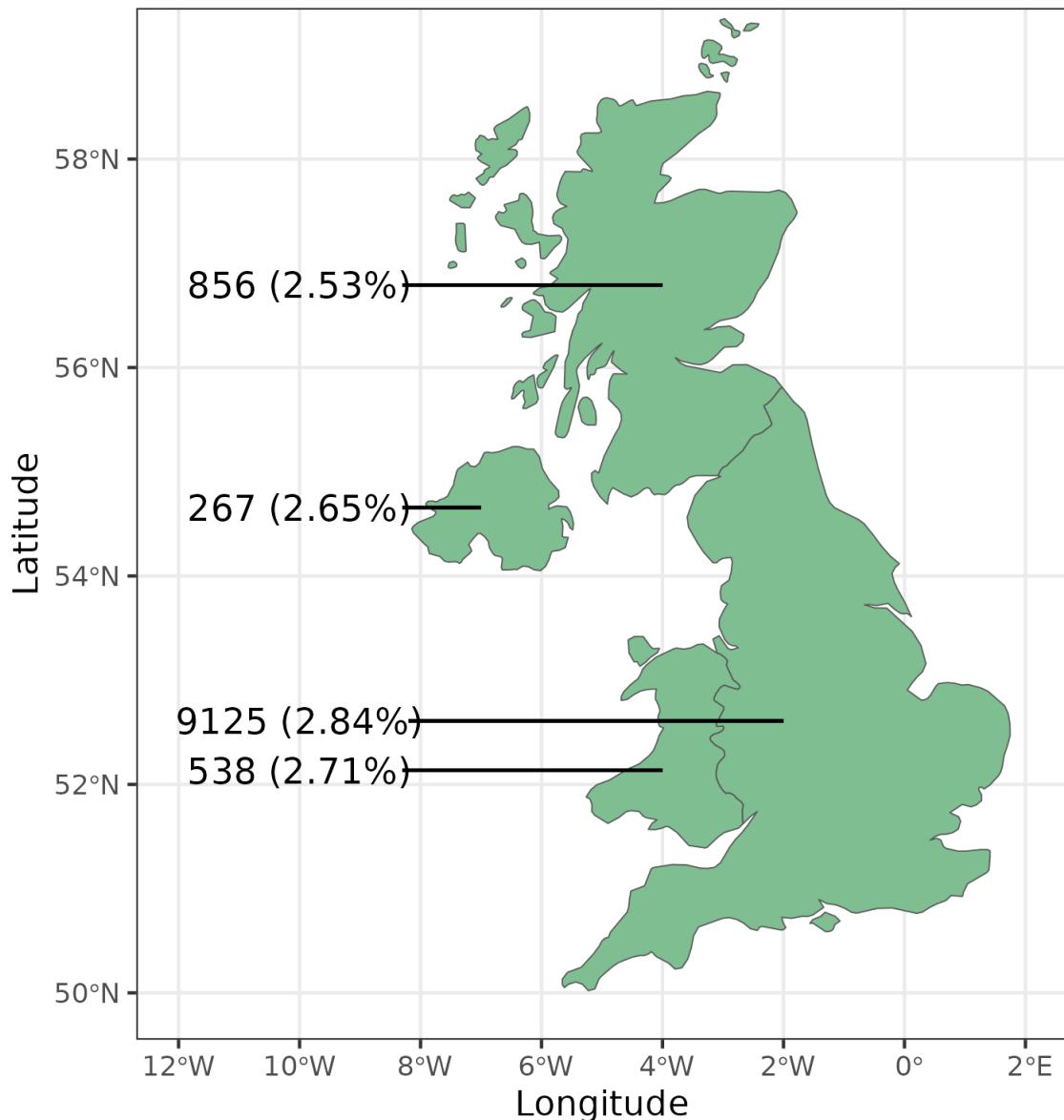


Figure 1.1: Pancreatic cancer incidence in the devolved nations 2017 to 2019

(SOR) ($C_{21}H_{16}ClF_3N_4O_3$), nab-paclitaxel (NAB)¹ and irinotecan (IRI) ($C_{33}H_{38}N_4O_6$). In addition, one standalone treatment, FOLFIRINOX, was included. FOLFIRINOX is a combination of oxaliplatin ($C_8H_{14}N_2O_4Pt$), irinotecan ($C_{33}H_{38}N_4O_6$), leucovorin ($C_{20}H_{23}N_7O_7$), and fluorouracil ($C_4H_3FN_2O_2$), which is currently the recommended first-line treatment for metastatic pancreatic cancer in the UK.

The National Institute for Health and Care Excellence (NICE) last updated their guidance on the diagnosis and management of pancreatic cancer in 2018 [National Institute for Health and Care Excellence, 2018]. The guidance for treatment is split for patients with locally advanced cancer and metastatic cancer.

1.2.1 Locally Advanced Pancreatic Cancer

Systemic combination chemotherapy is offered to patients who are well enough to tolerate it. For those who are not well enough to tolerate combination therapy, GEM monotherapy is offered. CAP should be considered as the radiosensitiser².

¹As nab-paclitaxel is a mixture of paclitaxel ($C_{47}H_{51}NO_{14}$) with albumin protein, it does not have a standard chemical formula

²A radiosensitiser is a treatment that makes cancer cells more susceptible to radiotherapy

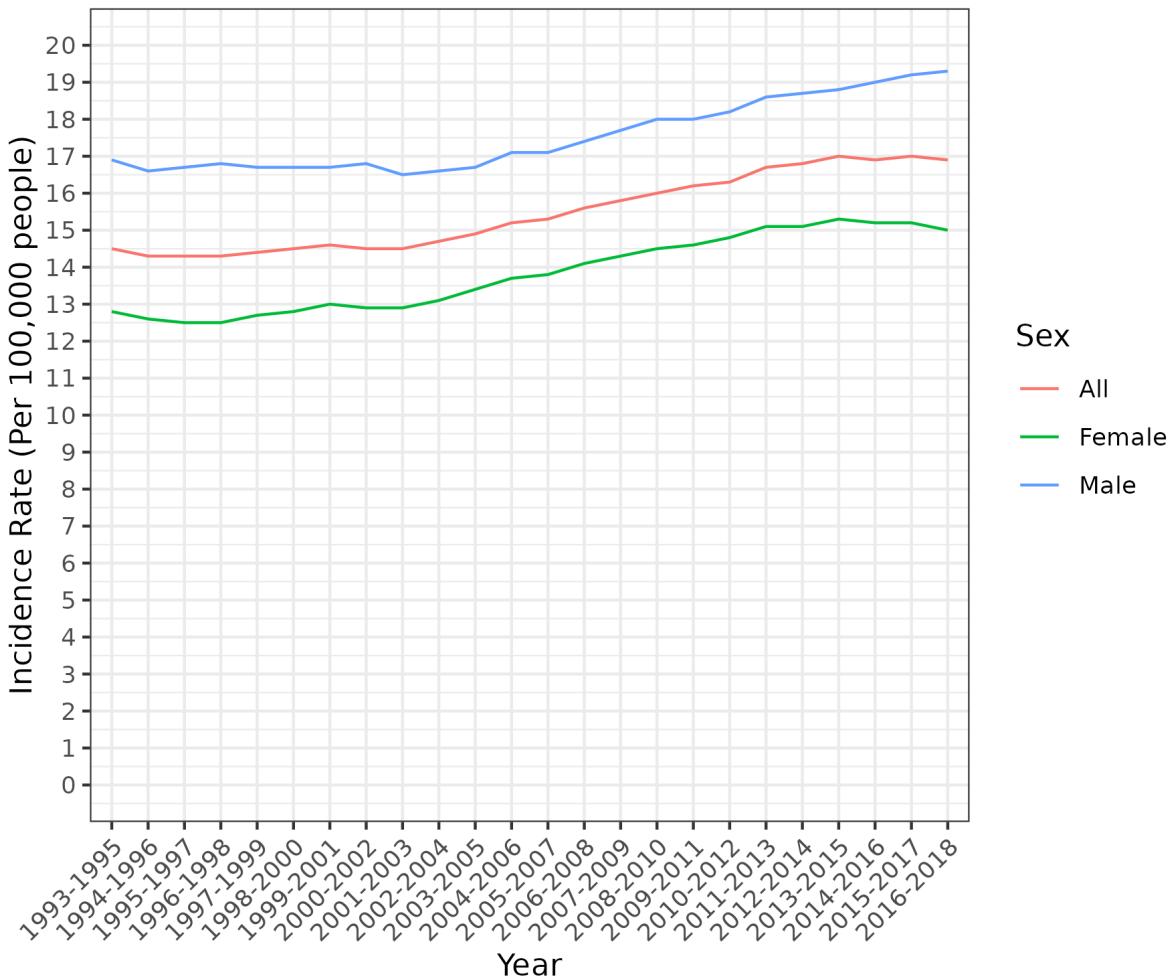


Figure 1.2: Pancreatic cancer incidence by sex, 1993 to 2018

1.2.2 Metastatic Pancreatic Cancer

In first line treatment, if a patient has an Eastern Cooperative Oncology Group (ECOG) performance status of 0 to 1 and is well enough, they are offered FOLFIRINOX. If the patient is not well enough to tolerate FOLFIRINOX, then GEM combination therapy is offered. Further, if the patient is not well enough to tolerate GEM combination therapy, then GEM monotherapy is offered.

In second-line treatment, oxaliplatin-based chemotherapy is offered to patients who did not receive first-line oxaliplatin. GEM combination therapy is offered in the second-line treatment case for patients who progressed after first-line FOLFIRINOX.

1.3 Project Aim

The primary objective was to assess the comparative efficacy of treatments for advanced/metastatic pancreatic cancer using Multilevel Network Meta-Regression (ML-NMR). The outcome of interest was Overall Survival (OS). As the focus of this project was not on a thorough literature review, but rather on the methodology, inclusion criteria for studies were not particularly strict. Studies had to be a phase II or phase III trial, and contain published Kaplan-Meier (KM) curves with numbers at risk. In addition, studies had to report the proportion of male patients on each treatment arm. Only the OS endpoint was considered.

The secondary aim was to corroborate the findings of current guidance and published literature. In particular, NICE TA476 [National Institute for Health and Care Excellence, 2017] described some un-

certainty in the comparison between GEM-NAB and GEM-CAP, so the Network Meta Analysis (NMA) also aimed to provide some clarification on the pairwise efficacy of GEM-NAB and GEM-CAP.

1.4 Dissertation Structure

Chapter 2 outlines some concepts in survival analysis that are important to this dissertation. The survival and hazard functions are introduced along with some key metrics. In particular, the median survival time and restricted mean survival time (RMST) are introduced. The KM estimator is derived using Maximum Likelihood Estimation (MLE), and an algorithm for reconstructing patient-level data from KM curves is presented. Finally, there is some discussion on parametric survival models, which the later chapters rely on. The process by which these models are fit is outlined, and the relevant R package for doing so.

Chapter 3 discusses the theory of NMAs. First, the idea of a network of evidence is outlined using graph theory. The standard NMA model is then discussed, looking at the relative effect of treatments in a network of evidence. The notion of different types of effects NMAs are introduced, and some discussion on types of variables considered in NMAs is presented. This is then expanded on to derive the ML-NMR model for general likelihoods, which is in turn expanded on to describe the ML-NMR model for survival outcomes. There is some discussion of the numerical integration techniques used when implementing the survival ML-NMR model. Following this, the outcome measures are introduced in the ML-NMR context, building on the initial definitions in Chapter 2. Finally, the model selection criteria and metrics are introduced.

Chapter 4 presents the clinical trials that were included in the NMA. A summary table presents the treatments, number of patients, and summary covariate information in each trial. There is some discussion on the similarity of the studies based on the data in the table. The discussion of similarity is aided by two forest plots presenting the variation of the median OS in each study. The KM curves for each study are also presented. Finally, the extrapolation plots of the parametric models fit to each treatment arm of each study are presented. The models are discussed to inform which models would be used as likelihoods in the NMA.

Chapter 6 presents the NMA itself. First, the network of evidence is presented along with some discussion. The model fitting and selection is then presented, with discussion on which model provided the best fit and was therefore selected. A table presenting the information criterion scores for each model is used to compare each model. The results are then presented. These include survival and hazard plots, along with forest plots of median OS and RMST.

Chapter 7 discusses the results of Chapter 6 and the interpretation thereof. The aims presented in Section 1.3 are addressed before some more general discussion and comments on potential future work. In addition, there is some discussion on how this NMA aligns with the good practice framework laid out by The Professional Society for Health Economics and Outcomes Research (ISPOR).

Appendix A presents some additional results from the NMA that were not required for drawing any conclusions, but may be of value to the interested reader. Appendix B details the individual ISPOR good practice questions, and the answers to them in the context of this NMA. Appendix C outlines the development of the PCNMA R package that was used for conducting this NMA. While not relevant to understanding any of the results, this appendix may be interesting to the more programming-minded reader.

SURVIVAL ANALYSIS BAKGROUND

2.1 Survival Functions

Given a homogeneous population of individuals, the time of death for each individual is drawn from a continuous random variable $T > 0$ with probability density function $f(t)$ and distribution function $F(t) = \int_0^t f(\tau)d\tau$. Survival analysis is concerned with estimating the distribution T from Time-To-Event (TTE) data. There are two functions central to survival analysis, the *survival function* and *hazard function*.

Definition 2.1.1: Survival Function

The **Survival Function** $S(t)$, gives the probability of an individual surviving longer than time t .

$$S(t) = P(T \geq t) = 1 - F(t) = \int_t^\infty f(\tau)d\tau \quad (2.1)$$

Definition 2.1.2: Hazard Function0

The **Hazard Function** gives the risk of death at time t , given that the individual has survived up to time t .

$$h(t) = -\frac{d}{dt} \log S(t) \quad (2.2)$$

Further to the survival and hazard function, the cumulative Hazard function, $H(t)$ (Definition 2.1) gives the total accumulated risk up to time t . By nature of being cumulative, $H(t)$ is a monotonically non-decreasing function.

Definition 2.1.3: Log-Cumulative Hazard Function

The Cumulative Hazard Function

$$H(t) = \int_0^t h(\tau)d\tau \quad (2.3)$$

$$= -\log(S(t)) \quad (2.4)$$

2.2 The Kaplan-Meier Estimator and Reconcstructing Patient-Level Data

2.2.1 The Kaplan-Meier Estimator

The KM estimator is a non-parametric method for estimating the survival function from event and censoring times. TTE data contains, at a bare minimum, a subject identifier, an event time, and a censoring indicator. The Censoring indicator is given in Equation 2.5. The time column of data gives the time that either the event (i.e, death due to the disease being investigated, in overall survival), or censoring (i.e a chemotherapy patient dies after being hit by a bus not their cancer) occurs.

$$c_i = \begin{cases} 1 & \text{If individual } i \text{ has an event} \\ 0 & \text{If individual } i \text{ is censored} \end{cases} \quad (2.5)$$

Let d_i and n_i be the number of events and total individuals at risk at the i^{th} timepoint, t_i . Define the discrete hazard rate h_i as the probability that individual experiences an event at time t_i . The survival rate is then defined as in Equation 2.6, and the likelihood function for the hazard function up to time t_i is given by Equation 2.7.

$$S(t) = \prod_{i:t_i \leq t} (1 - h_i) \quad (2.6)$$

$$\mathcal{L}(h_{j \leq i} | d_{j:j \leq i}, n_{j:j \leq i}) = \prod_{j=1}^i h_j^{d_j} (1 - h_j)^{n_j - d_j} \binom{n_j}{d_j} \quad (2.7)$$

The KM estimator can be derived by MLE of the discrete hazard function. By obtaining an maximum likelihood estimate of h_i , \hat{h}_i , and substituting it into Equation 2.6, the resulting estimate of the survival function, $\hat{S}(t)$, is the KM estimator. Taking the logarithm of both sides of Equation 2.7 gives Equation 2.8.

$$\log(\mathcal{L}) = \sum_{j=1}^i \left(d_j \log(h_j) + (n_j - d_j) \log(1 - h_j) + \log \binom{n_j}{d_j} \right) \quad (2.8)$$

By taking the derivative of Equation 2.8 with respect to h_i , and setting the resulting fraction equal to zero, \hat{h}_i is obtained as in Equation 2.11.

$$\frac{\partial \log(\mathcal{L})}{\partial h_i} = \frac{\partial}{\partial h_i} \left(\sum_{j=1}^i \left(d_j \log(h_j) + (n_j - d_j) \log(1 - h_j) + \log \binom{n_j}{d_j} \right) \right) \quad (2.9)$$

$$= \frac{d_i}{\hat{h}_i} - \frac{n_i - d_i}{1 - \hat{h}_i} \quad (2.10)$$

$$\Rightarrow \frac{\partial \log(\mathcal{L})}{\partial h_i} = 0 \Rightarrow \hat{h}_i = \frac{d_i}{n_i} \quad (2.11)$$

Definition 2.2.1: Kaplan-Meier Estimator

The **Kaplan-Meier Estimator**, $\hat{S}(t)$, estimates the survival probability that an individual survives longer than time t , and is given by Equation 2.12.

$$\hat{S}(t) = \prod_{i:t_i \leq t} \left(1 - \frac{d_i}{n_i} \right) \quad (2.12)$$

2.3 Reconstructing Survival Data from Published Curves

The ML-NMR method requires aggregate survival data to be in the form of the TTE data previously described. Often, the actual patient level event and censoring times/indicators are not published, and therefore must be reconstructed from the published KM curves, which are usually readily available. Indeed, one of the study selection criteria outlined in Section 1.3 was that studies must publish KM curves with numbers at risk. This was motivated by the method for reconstructing patient-level data from KM curves, the Guyot algorithm [Guyot et al., 2012] requiring numbers at risk.

2.4 Regression Models for Survival

The survival time of patients may be dependent on several explanatory variables such as age, sex, the presence of a genetic mutation, etc. We wish to incorporate these variables into our survival functions. There are two forms of models for survival data. Accelerated Failure Time (AFT) models, and Proportional-Hazards (PH) models).

2.4.1 Accelerated Failure Time Models

Let x be a vector of explanatory variables for each individual in a trial. The survival function can be extended to include this,

$$S(t, x) = S_0(t\Psi(x)).$$

Here, $S_0(t) = S(t, x = 0)$, i.e the survival function at baseline. We define the density and hazard functions accordingly,

$$\begin{aligned} f(t, x) &= f_0(t\Psi(x))\Psi(x) \\ h(t, x) &= h_0(t\Psi(x))\Psi(x). \end{aligned}$$

This is equivalent to defining a random variable T such that

$$T = T_0/\Psi(x).$$

Here, T_0 has survivor function S_0 . It is required that $\Psi(x) \geq 0$ and $\Psi(0) = 1$, leading to the natural choice

$$\Psi(x) = \exp(-\beta' x).$$

We can then write

$$\begin{aligned} T &= T_0/\Psi(x) \\ \implies E(T) &= E(T_0)/E(e^{-\beta' x}) \\ &= E(T_0)/e^{-\beta' x} \\ &= E(T_0)e^{\beta' x} \end{aligned}$$

In practice, we assume a distribution for T , and estimate parameters using maximum likelihood estimation.

2.4.2 Proportional Hazards Models

Let h_0 represent the hazard function for an individual at baseline. In addition, let \mathbf{x} be a vector of explanatory variables. The proportional hazards model, also known as the Cox model [Cox, 1972] is then given by

$$h(t, x) = \exp(\beta' x) h_0(t) \tag{2.13}$$

Consider the following definition.

Definition 2.4.1: Semi Parametric Model

A statistical model is a parameterised family of distributions $\{P_\theta : \theta \in \Theta\}$. For a parametric model, $\Theta \subseteq \mathbb{R}^k$ for $k \in \mathbb{N}$. Similarly, for a non-parametric model, $\Theta \subseteq V$, where V is some (possibly infinite) dimensional space V . A **Semi-parametric** model is a statistical model with both parametric and non-parametric components. For a semi-parametric model we have $\Theta \subseteq \mathbb{R}^k \times V$.

The Cox model is semi-parametric then β is of finite dimension and $h_0(t)$ is infinite-dimensional and does not need to be specified.

2.5 Key Survival Metrics

2.5.1 The Hazard Ratio

The Hazard Ratio (HR) follows from Equation 2.13. Consider two treatments, $i = 1, 2$, then $h_1(t, x) = \exp(\beta' x) h_0(t)$ and $h_2(t, x) = \exp(\beta' x) h_0(t)$. The HR is obtained as in Equation 2.14.

$$HR = \frac{h_1}{h_2} = \exp(\beta' x) \quad (2.14)$$

In practice, the HR is a useful endpoint in performing network meta-analyses on survival outcomes. However, in order to conduct a HR-based Network Meta Analysis (NMA), the proportional hazards assumption (PHA), must be satisfied. The PHA is the assumption that the HR remains constant throughout the observation period of a trial. It can be tested by visual-inspection of a log-cumulative hazards plot.

Definition 2.5.1: Cumulative Hazard Function

The **Cumulative Hazard Function**, $H(t)$ is given by

$$H(t) = \int_0^t h(t)dt = -\log(S(t))$$

By extension, the log-cumulative hazard function is given by $\log(-\log(S(t)))$. When plotting this for both arms of a clinical trial, if the curves remain roughly parallel, the PHA is likely satisfied, but if they curves cross, it indicates violation of the PHA.

2.5.2 Restricted Mean Survival Time

The RMST is an alternative measure to the (log) HR in NMAs. RMST is the mean survival time up to a pre-specified time. This measure can be thought of visually as the area under the survival curve. Definition 2.5.2 presents the formal definition.

Definition 2.5.2: RMST

For a survival function $S(t)$, the **RMST** for some pre-specified time $x > 0$,

$$RMST = \int_0^x S(t)dt$$

2.5.3 Median Survival

Median survival is simply the earliest timepoint at which 50% of patients have died.

2.6 Parametric Models for Survival Analysis

This section discusses the parametric models commonly used in Survival Analysis. In particular, the seven parametric models recommended by the NICE in Technical Support Document (TSD) 14 [Latimer, 2011]. All parametric model fitting for this project was performed in R using the **flexsurv** package [Jackson, 2016]. The first section outlines how the **flexsurv** package works.

2.6.1 Model Setup

The general model of a **flexsurv** survival model takes the form

$$f(t|\mu(\mathbf{z}), \alpha(\mathbf{z})). \quad (2.15)$$

Equation 2.15 gives the probability density for death at time $t \geq 0$. The *mean* or *location* of the distribution is given by $\mu = \alpha_0$. The remaining parameters, $\alpha^1 = (\alpha_1, \dots, \alpha_R)$ are called *ancillary* parameters.

Chapter 2 discussed AFT and PH models. Under the **flexsurv** framework, if the hazard function, $h(t) = \frac{f(t)}{S(t)}$, can be factorised as

$$h(t|\alpha, \mu(\mathbf{z})) = \mu(\mathbf{z})\mathbf{h}_0(\mathbf{t}|\alpha).$$

Then we have a PH model. On the other hand, an AFT model would be written as

$$S(t|\mu(\mathbf{z}), \alpha) = \mathbf{S}_0(\mu(\mathbf{z})\mathbf{t}/\alpha).$$

All parameters may depend on \mathbf{z} , a vector of covariates. This is done through the link-transformed linear models

$$\begin{aligned} g_0(\mu(\mathbf{z})) &= \gamma_0 + \beta_0^T \mathbf{z} \\ g_r(\alpha_r(\mathbf{z})) &= \gamma_r + \beta_r^T \mathbf{z} \end{aligned} \quad (2.16)$$

g is usually chosen to be $\log()$ if the parameter is positive, or the identity function if the parameter is unrestricted.

2.6.2 Fitting Models

Let $t_i, i \in \{1, \dots, n\}$, be a sample of times from n individuals. Define c_i as in Equation 2.5. Introduce s_i , which are delayed-entry times. This means for an individual i who is delayed-entry, the survival time is only observed conditionally on individual i having survived up to time s_i . $s_i = 0$ when there is no delayed-entry.

2.6.2.1 Right Censoring

In the case of right-censoring and nothing else, the likelihood for the parameters $\theta = \{\gamma, \beta\}$ required in Equation 2.16 is given by

$$l(\theta|\mathbf{t}, \mathbf{c}, \mathbf{s}) = \frac{\prod_{i:c_i=1} \mathbf{f}_i(\mathbf{t}_i) \prod_{i:c_i=0} \mathbf{S}_i(\mathbf{t}_i)}{\prod_i \mathbf{S}_i(s_i)} \quad (2.17)$$

2.6.2.2 Interval Censoring

In the case of interval-censoring, where the survival time is censored on (t_i^{\min}, t_i^{\max}) , the likelihood for $\theta = \{\gamma, \beta\}$ is

$$l(\theta|t^{\min}, t^{\max}, \mathbf{c}, \mathbf{s}) = \frac{\prod_{i:c_i=1} f_i(t_i) \prod_{i:c_i=0} (S_i(t_i^{\min}) - S_i(t_i^{\max}))}{\prod_i S_i(s_i)} \quad (2.18)$$

Maximum Likelihood Estimation is performed in R using the analytic derivatives of Equation 2.17 and/or Equation 2.18.

NETWORK META ANALYSIS THEORY

3.1 Building a Network of Evidence

Consider a set of N two-arm randomised-controlled trials (RCTs). In each trial $i \in 1, \dots, N$, the patients are randomised to receive a treatment A_i , or a placebo P_i . This can be represented as N graphs with two nodes, A_i and P_i , connected by an edge representing the trial comparing A_i and P_i . It is useful at this stage to recall the formal definition of an (undirected) graph.

Definition 3.1.1: Graph

A **Graph** is an ordered triple $G = (V, E, \varphi)$. Where V is a set of nodes, E is a set of edges, and $\varphi : E \rightarrow \{\{x, y\} | x, y \in V \text{ such that } x \neq y\}$ is an **incidence function** mapping every edge to a pair of vertices.

We can construct N graphs under the formal definition. Namely, for trial T_i , we have $G_i = (V_i, E_i, \varphi_i)$ where $V_i = \{A_i, P_i\}$, $E_i = \{T_i\}$ and $\varphi_i : E_i \rightarrow \{\{x, y\} | x, y \in V_i \text{ such that } x \neq y\}$. For construction of the graphs, we can drop the subscript on P_i , and take the placebo as a reference treatment. This is done under the assumption that the effect of placebo is constant across all trials. This is a strong assumption, and implications of this are discussed later. Under this assumption however, each V_i now contains a common element, P .

Let

$$V_{trts} = \bigcup_{i=1}^N V_i$$

$$E_{trials} = \bigcup_{i=1}^N E_i.$$

The incidence function is given in Equation 3.1.

$$\varphi : E_{trials} \rightarrow \{\{x, y\} | x, y \in V_{trts} \text{ such that } x = P\} \quad (3.1)$$

Then the ordered triple $G = (V_{trts}, E_{trials}, \varphi)$ is the network of evidence given by these two arm trials that forms the basis of an NMA. This process expands to trials that compare more than two treatments by weighting the edges by the number of trials making that particular comparison. Figure 3.1 illustrates combining treatments into a network of evidence.

In some scenarios, a head-to-head (HtH) clinical trial may be included in an NMA. Including a HtH trial involves special consideration, since these trials do not compare a treatment versus a placebo, but against another active treatment. From a graph-theoretic perspective, this can create a loop. A loop is an edge that connects a vertex to itself, and makes a graph non-simple.

3.2 Standard NMA Model

Let d_{ab} denote the relative effect of treatment b versus treatment a . Suppose we have summary outcomes y_{jk} of treatment k in study j . This summary outcome may be, for example, HRs, or RMST

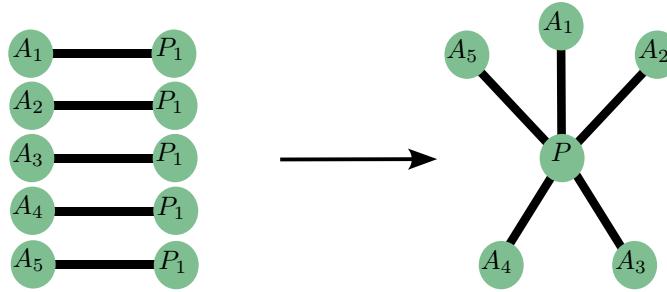


Figure 3.1: Visualisation of combining trials into a network of evidence

values. The standard NMA model is written as in Equation 3.2-Equation 3.3. In Equation 3.2, π_{Agg} is a suitable likelihood for the aggregate data, and θ_{jk} represents the expected summary outcome of treatment k in study j . The link function g serves to transform θ_{jk} onto the linear predictor scale. In Equation 3.3, μ_j and δ_{jk} are study-specific intercepts and study-specific relative effect of treatment k versus the reference treatment.

$$y_{jk} \sim \pi_{Agg}(\theta_{jk}) \quad (3.2)$$

$$g(\theta_{jk}) = \mu_j + \delta_{jk} \quad (3.3)$$

There are two types of NMAs: fixed effect (FE) and relative effects (RE) NMAs. In an FE NMA, $\delta_{jk} = d_{1k} = d_k$, with $d_1 = 0$. In an RE NMA, $\delta_{jk} \sim N(d_k, \tau^2)$ for the heterogeneity variance τ^2 , with $\delta_{j1} = d_1 = 0$.

The standard NMA model assumes that any effect modifiers, i.e covariates that alter the relative effect on a given scale of an active treatment versus control, are balanced across populations. While this can often be a valid assumption, methods such as Matching-Adjusted-Indirect-Comparisons (MAICs), Simulated Treatment Comparisons (STCs), and mlnmrags have sought to relax this assumption by using IPD from at least one of the studies in a population.

3.3 Multilevel Network Meta-Regression

The derivation in this chapter is based on the work of [Phillippo et al., 2024]. Under an NMA framework, there are J RCTs investigating a subset $K_j \subset K$ ($j = 1, \dots, J$) treatments. In this project, $|K_j| = 2 \forall j$. Depending on data availability, we may have individual-patient-data (IPD) for some studies, and only aggregate data for the remaining. This would be an ideal scenario, however it is not always the case.

Definition 3.3.1: General IPD Meta-Regression Model

Let y_{ijk} be the IPD outcome for individual $i = 1, \dots, N_{jk}$ in study j receiving treatment $k \in K_j$ given the likelihood distribution $\pi_{Ind}(\theta_{ijk})$.

$$\begin{aligned} y_{ijk} &\sim \pi_{Ind}(\theta_{ijk}) \\ g(\theta_{ijk}) &= \mu_j + x_{ijk}^T(\beta_1 + \beta_{2,k}) + \gamma_k \\ &= \eta_{jk}(x_{ijk}) \end{aligned}$$

Here, g links the likelihood parameter θ_{ijk} to $\eta_{jk}(x_{ijk})$. The μ_j are study-specific intercepts, and $\beta_1, \beta_{2,k}$ are regression coefficients for prognostic and effect-modifying covariates respectively. Additionally, the γ_k are individual-level treatment effects. For the reference treatment, $\beta_{2,1} = \gamma_1 = 0$.

It is clear to see how the model in Definition 3.3 extends Equations 3.2-3.3. Let $\xi = \{\mu_j, \beta_1, \beta_{2,k}, \gamma_k | \forall j, k\}$ be the parameter space. Using ξ , we can denote the individual conditional likelihood function by $L_{ijk|x}^{\text{Con}}(\xi; y_{ijk}, x_{ijk})$. The form of $L_{ijk|x}^{\text{Con}}$ depends on π_{Ind} , g , and η_{jk} .

By integrating the individual conditional likelihood over the joint covariate distribution f_{jk} , we obtain Equation 3.4, the individual marginal likelihood function.

$$L_{ijk}^{\text{Mar}}(\xi; y_{ijk}) = \int_{\mathcal{X}} L_{ijk|x}^{\text{Con}}(\xi; y_{ijk}, x) f_{jk}(x) dx \quad (3.4)$$

It is clear from Equation 3.4 that L_{ijk}^{Mar} does not depend on x . Let i be an individual on treatment k in study j with outcome y_{ijk} . If we do not know the covariate vector for i , x_{ijk} , but we do know f_{jk} , then we know that the likelihood contribution of i is given by Equation 3.4.

It is likely that a closed-form of Equation 3.4 does not exist. We can therefore take a set of N integration points, \hat{x} from f_{jk} , and approximate the integral using Monte-Carlo methods, giving Equation 3.5.

$$L_{ijk}^{\text{Mar}}(\xi; y_{ijk}) \approx \frac{1}{N} \sum_{\hat{x}} L_{ijk|x}^{\text{Con}}(\xi; y_{ijk}, \hat{x}) \quad (3.5)$$

Consider now a summary outcome y_{jk} aggregated over all individuals on treatment k in study j . Each individual i in on treatment k in study j contributes to the aggregate likelihood. Let y_{ijk} denote the observed value of this summary measure for individual i . The aggregate marginal likelihood function is then the product of these y_{ijk} up to a normalising constant, as in Equation 3.6.

$$L_{jk}^{\text{Mar}} \propto \prod_{i=1}^{N_{jk}} L_{ijk}^{\text{Mar}}(\xi; y_{ijk}) \quad (3.6)$$

The full, general, ML-NMR model is then given by combining the individual and aggregate level components, as in Definition 3.3.

Definition 3.3.2: General ML-NMR Model

Individual:

$$L_{ijk|x}^{\text{Con}}(\xi; y_{ijk}, x_{ijk}) = \pi_{\text{Ind}}(y_{ijk} | \theta_{ijk}) \quad (3.7)$$

$$g(\theta_{ijk}) = \eta_{jk}(x_{ijk}) = \mu_j + x_{ijk}^T(\beta_1 + \beta_{2,k}) + \gamma_k \quad (3.8)$$

Aggregate:

$$L_{ijk}^{\text{Mar}}(\xi; y_{ijk}) = \int_{\mathcal{X}} L_{ijk|x}^{\text{Con}}(\xi; y_{ijk}, x) f_{jk}(x) dx \quad (3.9)$$

$$L_{jk}^{\text{Mar}} \propto \prod_{i=1}^{N_{jk}} L_{ijk}^{\text{Mar}}(\xi; y_{ijk}) \quad (3.10)$$

Under a Bayesian framework, priors are placed on μ_j , β_1 , $\beta_{2,l}$, and γ_k . Uninformative priors were used.

3.4 Survival ML-NMR

Each study reports a pair $y_{ijk} = \{t_{ijk}, c_{ijk}\}$ consisting of outcome times t_{ijk} and censoring indicators c_{ijk} , either from IPD or reconstructed IPD. For IPD studies, the covariates x_{ijk} will naturally be available, but for aggregate studies (those for which pseudo-IPD has been re-created), only the joint covariate distribution of covariates at baseline, denoted f_{jk} .

With some covariates, the pseudo-IPD can include a recreated covariate column for each individual. For example, if the proportion of male patients is given for a summary study, pseudo-patients can be randomly assigned to be male or female such that the number matches that reported in the trial. This is not possible with covariates such as age, where a distribution is usually reported. The proportion of male patients is somewhat unique in this case. It would likely not be statistically sound to do this with a variable such as the site of any metastases. Despite this being a count variable in the same way, the site of metastases may have more influence on survival times than the sex of a patient, and therefore

randomly assigning patients to have a given metastases based on the aggregate data may not align with what the original data found.

The censoring indicator for patient i in study j on treatment k is defined as in Equation 2.5. In practice, the censoring indicator can be the other way round- i.e 1 denotes censoring instead of an event, as in Equation 2.5. Therefore, when cleaning data for this dissertation, manual reversing of the censoring indicator was conducted to ensure all data used the same definition.

Let $S_{jk}(t|\mathbf{x})$ and $h_{jk}(t|x)$ be the survival and hazard functions conditional on the covariates x . Then the individual conditional likelihood contributions for each time t_{ijk} in the IPD studies are given by

$$L_{ijk|x}^{Con}(\zeta; t_{ijk}, c_{ijk}, x_{ijk}) = S_{jk}(t_{ijk}|x_{ijk})h_{jk}(t_{ijk}|x_{ijk})^{c_{ijk}} \quad (3.11)$$

The forms of S and h depend on the specific survival models chosen. Starting from Equation 3.4, the marginal likelihood equations for each event/censoring time in the aggregate data studies can be derived. Substituting $y_{ijk} = \{t_{ijk}, c_{ijk}\}$

$$L_{ijk}^{\text{Mar}}(\xi; y_{ijk}) = \int_{\mathbf{x}} L_{ijk|x}^{\text{Con}}(\xi; y_{ijk}, x) f_{jk}(x) dx \quad (3.12)$$

$$L_{ijk}^{\text{Mar}}(\xi; t_{ijk}, c_{ijk}) = \int_{\mathbf{x}} L_{ijk|x}^{\text{Con}}(\xi; t_{ijk}, c_{ijk}, x) f_{jk}(x) dx \quad (3.13)$$

$$= \int_{\mathbf{x}} S_{jk}(t_{ijk}|x_{ijk})h_{jk}(t_{ijk}|x_{ijk})^{c_{ijk}} f_{jk}(x) dx \quad (3.14)$$

As with Equation 3.5, Equation 3.14 can be evaluated with quasi-Monte Carlo integration to obtain Equation 3.15.

$$L_{ijk}^{\text{Mar}}(\xi; t_{ijk}, c_{ijk}) = \frac{1}{N} \sum_{\hat{x}} S_{jk}(t_{ijk}|\hat{x})h_{jk}(t_{ijk}|\hat{x})^{c_{ijk}} \quad (3.15)$$

Quasi-Monte-Carlo (QMC) is a method for efficient numerical integration. Equation 3.16 gives the form of a Monte-Carlo integration problem for a real integrable function f over the s -dimensional hypercube $I^s = [0, 1]^s$. QMC differs from standard Monte-Carlo (MC) integration in the way the x_i are chosen for which f is evaluated at.

$$\int_{[0,1]^s} f(u) du \approx \frac{1}{N} \sum_{i=1}^N f(x_i) \quad (3.16)$$

Regular MC integration uses random numbers¹, whereas QMC uses a low-discrepancy sequence, which is a sequence such that for all N , the set $P = \{x_1, \dots, x_N\}$ has low discrepancy, as defined in Definition 3.4.

Definition 3.4.1: Low-Discrepancy Sequence

Consider the set $P = \{x_1, \dots, x_N\}$. The **discrepancy** of P , $D_N(P)$ is defined as in Equation 3.17.

$$D_N(P) = \sup_{B \in J} \left| \frac{A(B; P)}{N} - \lambda_s(B) \right| \quad (3.17)$$

Here, $A(B; P)$ is the number of points in P that fall into B . λ_s is the s -dimension Lebesgue measure, and J is the set of s -dimensional intervals as in Equation 3.18, where $0 \leq a_i < b_i \leq 1$.

$$\prod_{i=1}^s [a_i, b_i) = \{x \in \mathbb{R}^s | a_i \leq x_i < b_i\} \quad (3.18)$$

In particular, `multinma` uses a Sobol sequence [Sobol', 1967] to sample \hat{x}_{jk} from the covariate distribution f_{jk} [Phillippo et al., 2020]. For each covariate in the analysis, the Sobol sequence generates points in $I^s = [0, 1]^s$ with a dimension for each covariate. The Smirnov transform (Definition 3.4) is then used to transform the points generated by the Sobol sequence to the required distribution.

¹Strictly speaking, these are pseudo-random numbers when considering the implementation of this method.

Definition 3.4.2: Smirnov Transformation

Let $X \in \mathbb{R}$ be any random variable. Then the random variable $F_X^{-1}(U)$ has the same distribution as X . Here, F_X^{-1} is the generalised inverse of the cumulative distribution function F_X of X , as defined in Equation 3.19 for all $p \in [0, 1]$. Further, $U \sim \text{Uniform}[0, 1]$.

$$F^{-1}(p) = \inf\{x \in \mathbb{R} | F(x) \geq p\} \quad (3.19)$$

3.5 Population-average estimates

Recall d_{ab} is the relative effect of treatment b versus treatment a. Let $d_{ab(P)}$ be the population-average relative effect of b versus a in population P. $d_{ab(P)}$ can be calculated as in Equation 3.21.

$$d_{ab(P)} = \int_{\mathfrak{X}} (\eta_{(P)b}(x) - \eta_{(P)a}(x)) f_{(P)}(x) dx \quad (3.20)$$

$$= \gamma_b - \gamma_a + \bar{x}_{(P)}^T (\beta_{2,b} - \beta_{2,a}) \quad (3.21)$$

For this dissertation, The estimates considered for results were the RMST and median OS. The survival and hazard functions were also analysed but were not primary outcomes for the NMA.

3.5.1 Survival function

Let $\bar{S}_{(P)k}(t)$ be the population-average marginal survival probability of treatment k in population P at time t. $\bar{S}_{(P)k}(t)$ is obtained by integrating $S_{(P)k}(t|x)$ over $f_{(P)}(x)$, as in Equation 3.22.

$$\bar{S}_{(P)k}(t) = \int_{\mathfrak{X}} S_{(P)k}(t|x) f_{(P)}(x) dx \quad (3.22)$$

3.5.2 Hazard function

The population-average marginal hazard function and cumulative hazard function are given by Equation 3.23, and Equation 3.24, respectively.

$$\bar{h}_{(P)k}(t) = \frac{\int_{\mathfrak{X}} S_{(P)k}(t|x) h_{(P)k}(t|x) f_{(P)k}(x)}{\bar{S}_{(P)k}(t)} \quad (3.23)$$

$$\bar{H}_{(P)k}(t) = -\log(\bar{S}_{(P)k}(t)) \quad (3.24)$$

3.5.3 RMST

Let x be some time horizon. The population-average marginal RMST follows from Definition 2.5.2.

$$RMST_{(P)k}(x) = \int_0^x \bar{S}_{(P)k}(t) dt. \quad (3.25)$$

3.5.4 Median OS

In general, the $\alpha\%$ quantile is obtained by solving

$$\bar{S}_{(P)k}(t) = 1 - \alpha. \quad (3.26)$$

Since the median OS is a special case of this with $\alpha = \frac{1}{2}$, the population-average marginal median OS, m is estimated by Equation 3.27.

$$\bar{S}_{(P)k}(m) = \frac{1}{2} \quad (3.27)$$

3.6 Model Selection and Convergence

Model selection was based primarily on the Leave-One-Out Information Criterion (LOOIC) score (Defintion 3.6). The LOOIC was calculated using the `loo` R package [Vehtari et al., 2024]. The model with the lowest LOOIC score was selected as the best fitting model. The Deviation Information Criterion (DIC) score (Definition 3.6) was used a secondary selection criterion. Again, models with lower DIC indicated better fit.

Definition 3.6.1: Leave-One-Out Information Criterion

Begin by defining the expected log pointwise predictive density (elpd) as in Equation 3.28 for data y ,

$$\text{elpd}_{\text{loo}} = \sum_{i=1}^n \log(p(y_i|y_{-i})) \quad (3.28)$$

Here, $p(y_i|y_{-i})$ is the LOO predictive density given the data y with the i^{th} data point removed.

$$p(y_i|y_{-i}) = \int p(y_i|\theta)p(\theta|y_{-i})d\theta \quad (3.29)$$

Define the **Leave One Out Information Criterion (LOOIC)** score as in Equation 3.30.

$$\text{LOOIC} = -2\text{elpd}_{\text{loo}} \quad (3.30)$$

Definition 3.6.2: Deviation Information Criterion (DIC)

Let y be the data on which a model was fitted. Further, let θ and $p(y|\theta)$ be the unknown parameters and likelihood function respectively. Let $E(\theta) = \bar{\theta}$ be the expectationi of θ . The deviance is defined as in Equation 3.31.

$$D(\theta) = -2 \log(p(y|\theta)) + C \quad (3.31)$$

Here, C is an unknown constant. C cancels in all model comparisons so is of no importance, but is included here for completeness. Further, the effective number of parameters is given by Equation 3.32.

$$p_D = D(\bar{\theta}) - D(\bar{\theta}) \quad (3.32)$$

The **Deviation Information Criterion (DIC)** is then given by Equation 3.33.

$$DIC = p_D + D(\bar{\theta}) \quad (3.33)$$

The LOOIC score was preferred over the DIC score as the DIC score is based on a point estimate, and this NMA was performed in a Bayesian framework. As the LOOIC score is uses the posterior distribution, it is a fully Bayesian statistic. Traditionally, the DIC score has been used due to the extra computational steps required for the LOOIC score. However, [Vehtari et al., 2017] impleted Pareto-Smoothed Importance Sampling (PSIS) in the R package `loo` [Vehtari et al., 2024], which enables easy, and stable, computation of the LOOIC score from the NMA models.

CHAPTER
FOUR

INCLUDED STUDIES

Table 4.1 presents the studies used in this NMA. In total, there were seven studies comparing GEM with one of six combination therapies. The studies were comparable in terms of median age and proportion male. The [Cunningham et al., 2009] study, compared GEM and GEM-CAP. The [Goldstein et al., 2015] study compared GEM and GEM-NAB. The [Gonçalves et al., 2012] study compared GEM and GEM-SOR. The [Kindler et al., 2011] and [Spano et al., 2008] studies compared GEM and GEM-AXI. The [Oettle et al., 2005] study compared GEM and GEM-PEM. The [Rocha Lima et al., 2004] study compared GEM and GEM-IRI. All studies except Spano were phase III trials, however Spano was included as the OS data was quite mature. The [Conroy et al., 2011] study was the only study to not compare GEM with a combination therapy; comparing GEM and FOLFIRINOX in a phase 2-3 trial.

Figure 4.1 and Figure 4.2 present forest plots of the median OS of the GEM arm in each study, and the comparator arm in each study, respectively. In particular, the GEM arms in the Conroy, Cunningham, Oettle, Goldstein, and Rocha Lima studies were similar, while the Spano study was noted for having a large 95% CI. This was to be expected given there were only 34 patients in the GEM arm. For the comparator arms, there was more variation in the reported median OS. Large 95% CIs were present for the Spano and Goncalves studies, again due to the comparatively low number of patients in these studies. The efficacy of FOLFIRINOX is clear from Figure 4.2. The median OS of FOLFIRINOX is above the upper bound of 95% credible interval of all comparators.

Study	Treatment	N	Median Age	Proportion Male	Median OS (Months)
Conroy 2011	GEM	171	61.0 (34, 75)	0.620	6.8 (5.5, 7.6)
Conroy 2011	FOLFIRINOX	171	61.0 (25, 76)	0.614	11.1 (9.0, 13.1)
Cunningham 2009	GEM	266	62.0 (26, 83)	0.580	6.2 (5.5, 7.2)
Cunningham 2009	GEM-CAP	267	62.0 (37, 82)	0.570	7.1 (6.2, 7.8)
Goldstein 2015	GEM	430	63.0 (32, 88)	0.600	6.6 (6.0, 7.2)
Goldstein 2015	GEM-NAB	431	62.0 (27, 86)	0.570	8.7 (7.9, 9.7)
Goncalves 2012	GEM	52	64.0 (40, 82)	0.620	9.2 (7.7, 11.6)
Goncalves 2012	GEM-SOR	52	61.0 (42, 85)	0.580	8.0 (6.0, 10.8)
Kindler 2011	GEM	316	61.0 (35, 89)	0.590	8.3 (6.9, 10.3)
Kindler 2011	GEM-AXI	314	61.0 (34, 84)	0.610	8.5 (6.9, 9.5)
Oettle 2005	GEM	282	63.0 (28, 82)	0.535	6.3 (5.4, 6.9)
Oettle 2005	GEM-PEM	283	63.0 (27, 82)	0.604	6.2 (5.4, 6.9)
Rocha Lima 2004	GEM	180	60.2 (32, 83)	0.533	6.6 (5.2, 7.8)
Rocha Lima 2004	GEM-IRI	180	63.2 (39, 81)	0.572	6.3 (4.7, 7.5)
Spano 2008	GEM	34	61.0 (36, 78)	0.470	5.6 (3.9, 8.8)
Spano 2008	GEM-AXI	69	65.0 (44, 81)	0.510	6.9 (5.3, 10.1)

Table 4.1: Included studies with summary statistics

Figure 4.3 presents the KM curves for each treatment arm in each study. It was clear that the PHA would need to be relaxed when fitting NMA models from the shape of the curves in Figure 4.3, due to the amount of crossing. Most of the studies had mature data, however the Kindler and Goncalves studies were noted for only dropping to an OS of around 0.25 at the end of the observation period. Conroy, and Goldstein were the only studies that showed a clearly higher OS for the comparator compared to

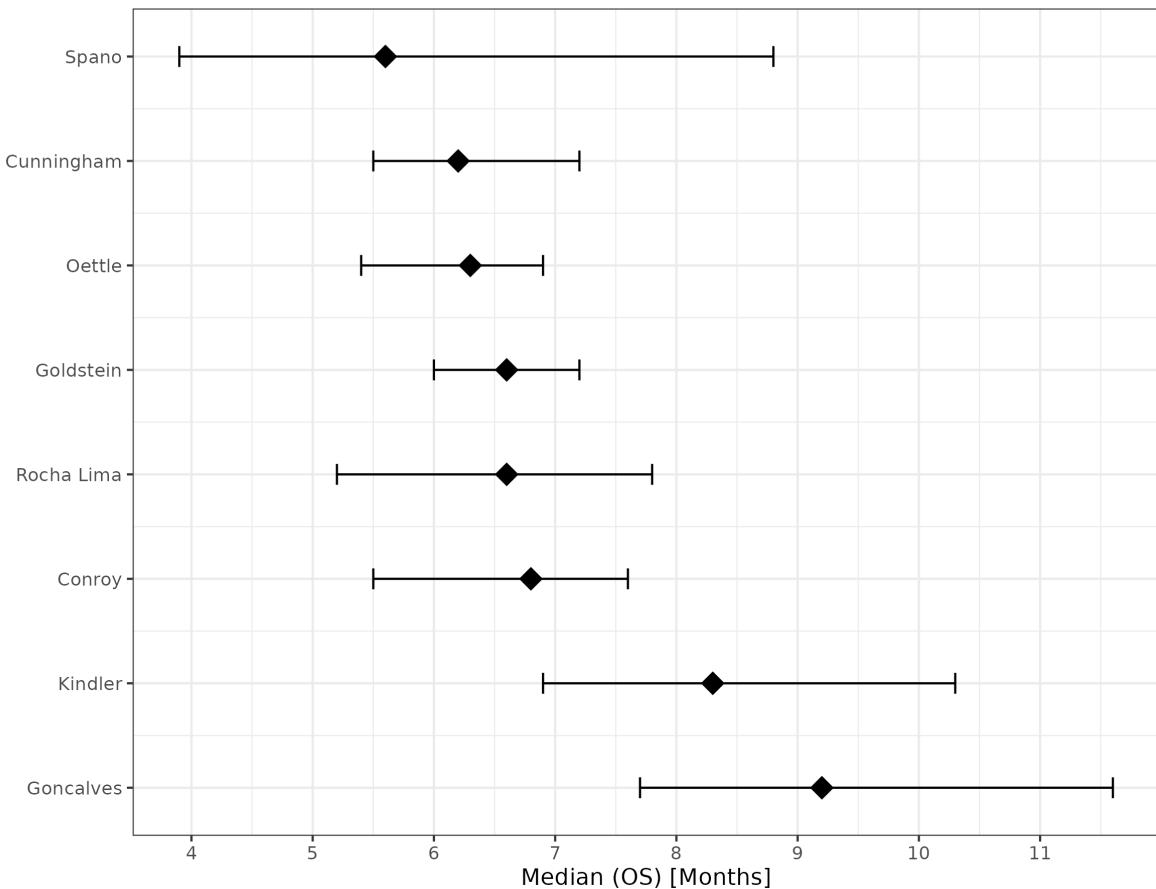


Figure 4.1: Forest plot for median OS of GEM in each study

GEM. The curves in the Cunningham, Goncalves, and Spano studies showed slight improvement in the comparator compared to GEM. The curves in the Kindler, Oettle, and Rocha Lima studies showed little to no benefit between the comparator and GEM curves.

To assess the PHA more formally, the log-cumulative hazard plot was used. If the curves cross, this indicates violation of the PHA. Figure 4.4 presents the log-cumulative hazard function for each study. None of the studies have non-crossing log-cumulative hazard curves. In particular, the Kindler and Oettle studies had log-cumulative hazard curves that were almost identical.

4.1 Study Eligibility Criteria

4.1.1 Eligibility by study

Patients in the Conroy study had to be aged 18 or over with cytologically confirmed, previously untreated, metastatic pancreatic adenocarcinoma. Patients had to have an ECOG score of 0 or 1. Patients were required to have bone marrow such that granulocyte count was $\geq 1500/mm^3$ and platelet count was $\geq 100,000/mm^3$. In addition to bone marrow, the renal and liver function had to be adequate. For the liver function, patients had to have bilirubin ≤ 1.5 times the upper limit of the normal range.

The Conroy study excluded patients who were aged 76 or over. In addition, endocrine or acinar pancreatic carcinoma, previous radiotherapy for measurable lesions, cerebral metastases, a history of another major cancer, active infection, chronic diarrhea, a clinically significant history of cardiac disease, and pregnancy or breast-feeding were all exclusion criteria.

The eligibility criteria in the Cunningham study were similar to Conroy, however patients with locally

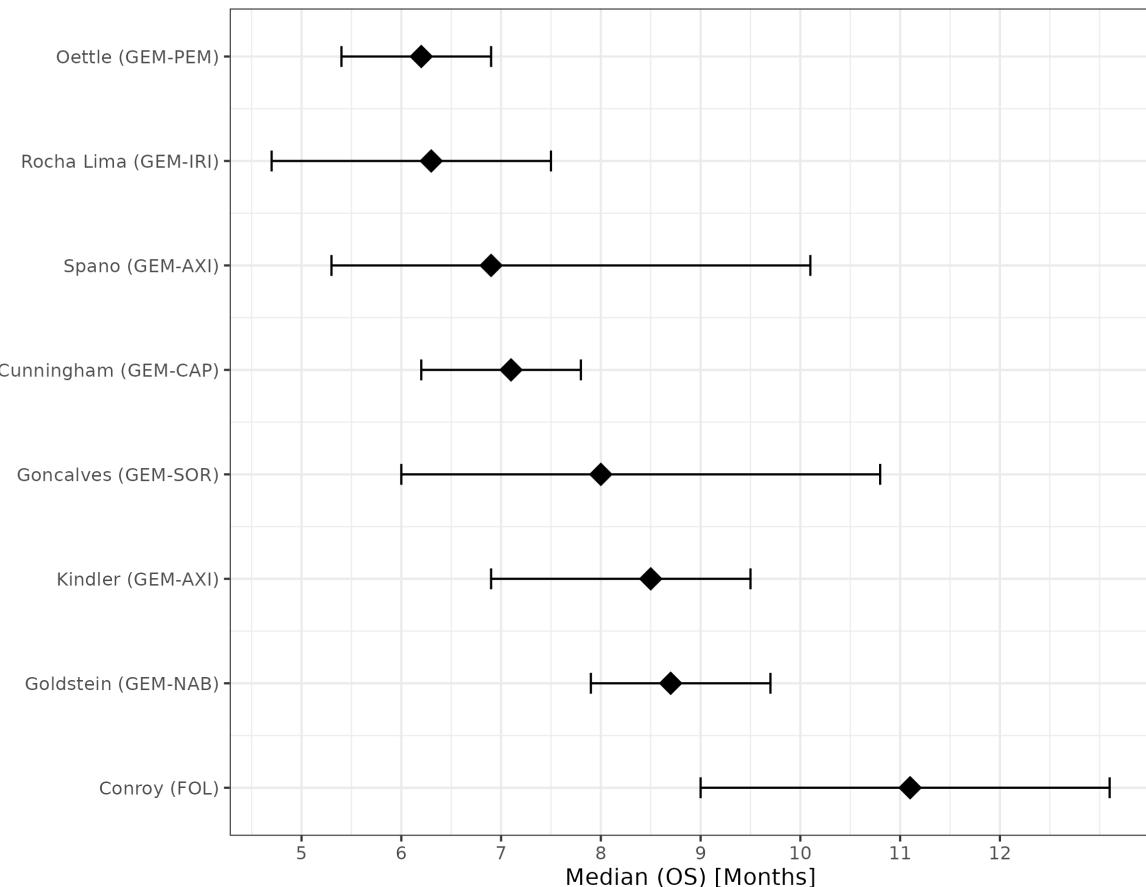


Figure 4.2: Forest plot for median OS of the comparator in each study

advanced disease were eligible in Cunningham. Patients could not have had previous chemotherapy or investigative treatment. Adequate bone marrow was required. There were no exclusion criteria.

The Goldstein study had eligibility in terms of the type of cancer closer related to Conroy than Cunningham. Eligible patients had metastatic cancer. In terms of the liver function, the requirement was a bilirubin level below or equivalent to the upper limit of the normal range. In addition, eligible patients had an absolute neutrophil count of $\geq 1.5 \times 10^9/L$ and a hemoglobin level $\geq 9g/dl$. The prior treatment eligibility was more relaxed in the Goldstein study. Treatment with either gemcitabine or fluorouracil as a radiosensitizer was allowed if it had been at least six months before randomisation.

As in Conroy, there were exclusion criteria in the Goldstein study. Previous chemotherapy for metastatic disease was an exclusion criterion. In addition, patients with islet cell neoplasms or locally advanced adenocarcinoma were also excluded. Patients who received cytotoxic doses of any systemic chemotherapy, including gemcitabine, in the adjuvant setting were excluded.

The eligibility criteria for the Goncalves study closer aligned with the Cunningham study. Patients with either locally advanced or metastatic cancer were eligible. An ECOG score between 0 and 2 was required, and patients had to be older than 18. In addition, the granulocyte and platelet counts were $> 1.5 \times 10^9/l$ and $> 100 \times 10^9/l$ respectively. Adequate bone marrow, liver, and renal function were required. The total bilirubin was required to be < 1.5 times the upper limit of the normal range.

Exclusion criteria for the Goncalves study were brain metastases, intestinal obstruction, a history of inflammatory bowel disease or extended small bowel resection. Patients could not have had any major surgery or radiotherapy within 28 days of randomisation.

In the Kindler study, patients had to be at least 18, and could have either locally advanced or metastatic pancreatic cancer. The ECOG score needed to be either 0 or 1. Adequate bone marrow,

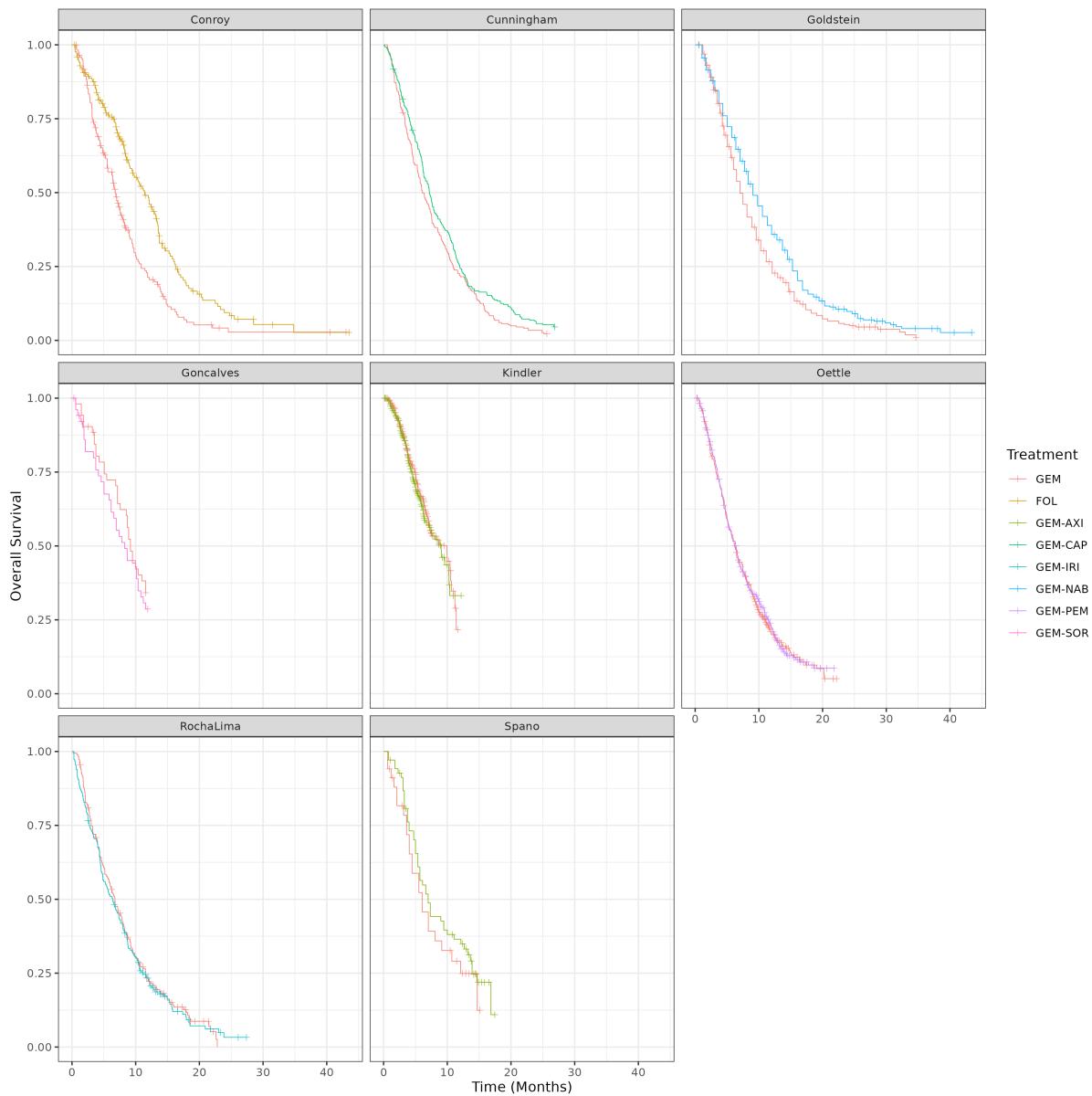


Figure 4.3: KM curves for each study

renal, and liver function was required, but no strict criteria was given. Adjuvant therapy was allowed provided it did not contain GEM, and at least four weeks had passed since the last dose.

Patients were excluded in the Kindler study if they had previous systemic chemotherapy for locally advanced or metastatic disease. In addition recent haemoptysis, myocardial infarction, symptomatic congestive heart failure, cerebrovascular accident, transient ischaemic attack, deepvein thrombosis, or pulmonary embolism in the past 12 months, peptic ulcer disease needing treatment in the past 6 months, active seizures or gastrointestinal bleeding were all exclusion criteria.

The eligibility criteria for the Oettle study were patients aged at least 18 with either locally advanced or metastatic pancreatic cancer. Prior radiotherapy was allowed if it was completed at least four weeks before entry into the Oettle study. ECOG scores between 0 and 2 were acceptable. Adequate bone marrow, renal, and liver function was required, but no performance criteria given. The Kindler study also required patients to have a life expectancy of at least 12 weeks.

Similar to the Goncalves study, patients with brain metastases were excluded. In addition, patients with significant weight loss were also excluded. This was defined as the loss of > 10% of bodyweight in

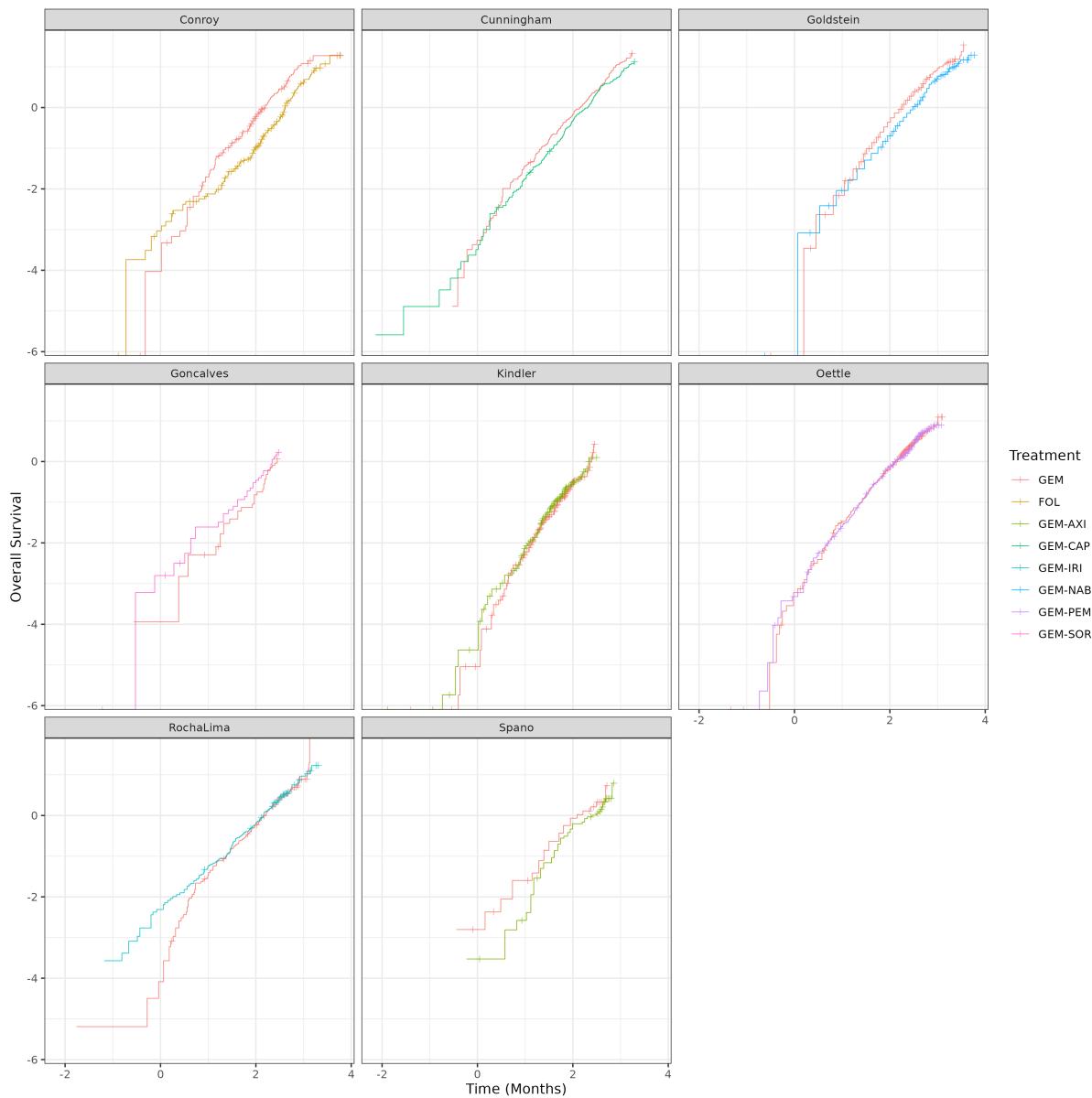


Figure 4.4: Log-cumulative hazard plot for each study

the previous six weeks. Patients who were unable to interrupt non-steroidal anti-inflammatory drugs for a 5- to 8-day period around PEM administration, or those unable to take folic acid or vitamin B_{12} were also ineligible.

The Rocha Lima study allowed for patients aged at least 18 with locally advanced or metastatic pancreatic cancer. The required ECOG score was between 0 and 2. The absolute neutrophil count, platelet count, and bilirubin levels were $\geq 1500/\mu L$, $\geq 100,00/\mu L$ and ≤ 1.5 times the upper limit of normal, respectively. These were in line with other studies included.

Patients were excluded from the Rocha Lima study if they had received prior systemic therapy either in an adjuvant setting or for the treatment of advanced pancreatic cancer. Patients could not be pregnant or breastfeeding, have active inflammatory bowel disease, significant bowel obstruction, chronic diarrhea, known brain disease, or myocardial infarction within the previous six months, uncontrollable high blood pressure, unstable angina, congestive heart failure, uncontrolled cardiac arrhythmia, HIV/AIDS or psychiatric illness that prevented the patient giving informed consent.

In the Spano study, patients who were aged 18 or older with locally advanced or metastatic pancreatic

cancer were eligible. An ECOG score between 0 and 2 was required. The absolute neutrophil count, platelet count, and hemoglobin levels were $\geq 1500/\mu\text{L}$, $\geq 100,00/\mu\text{L}$ and $\geq 90\text{g/L}$, respectively.

Patients were excluded from Spano if they received prior treatment for metastatic disease or treatment with GEM. Pregnancy or breast feeding, prior cerebrovascular accident, major surgery within the previous 4 weeks, brain metastases, active second malignancy, uncontrolled intercurrent illness, urine protein of 500mg or more in a 24 hour period, or ongoing uncontrolled hypertension were all exclusion criteria.

4.1.2 Analysis of eligibility

There was a lot of crossover in the eligibility criteria across the included studies. The two main differences is that the Conroy and Goldstein studies did not include patients with locally advanced disease, and that the Conroy study had an upper age limit for eligible patients. Based on this, it was deemed to be likely that the the populations were homogeneous. Without IPD it is impossible to fully assess homogeneity, but assuming the homogeneity based on the eligibility criteria, and summary statistics presented in Table 4.1 was deemed reasonable.

4.2 Covariates

No IPD was available for any study in the network. For the ML-NMR model to fit, at least one study needs to have IPD. To deal with this, the sex of patients in the Golstein study was simulated based on the reported proportion of male patients in each treatment arm. For the GEM and NAB treatment arms, 60% and 57% of patients were assigned to be male, respectively.

4.3 Parametric Model Fitting

Figure 4.5 to Figure 4.12 present the extrapolation plots for each treatment arm in each study. The data was mature in all studies except the Kindler and Goncalves studies, which meant there was more variation in the survival models for treatments in these populations. The exponential model was noted for presenting poor visual fit in both treatments across both of these studies. For this reason, and due to the NMA not assuming the PHA held, the exponential model was left out of the NMA. Based on these plots, all other models were deemed to be appropriate for inclusion in the NMA.

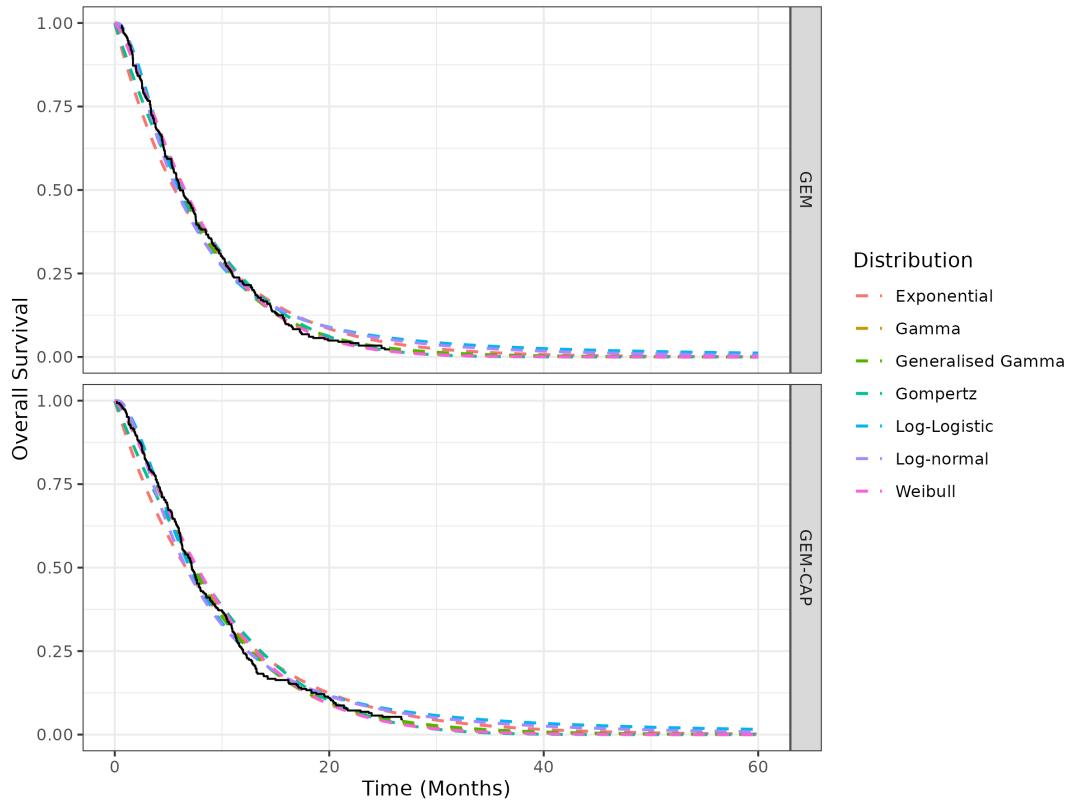


Figure 4.5: Cunningham (2009) parametric model extrapolations

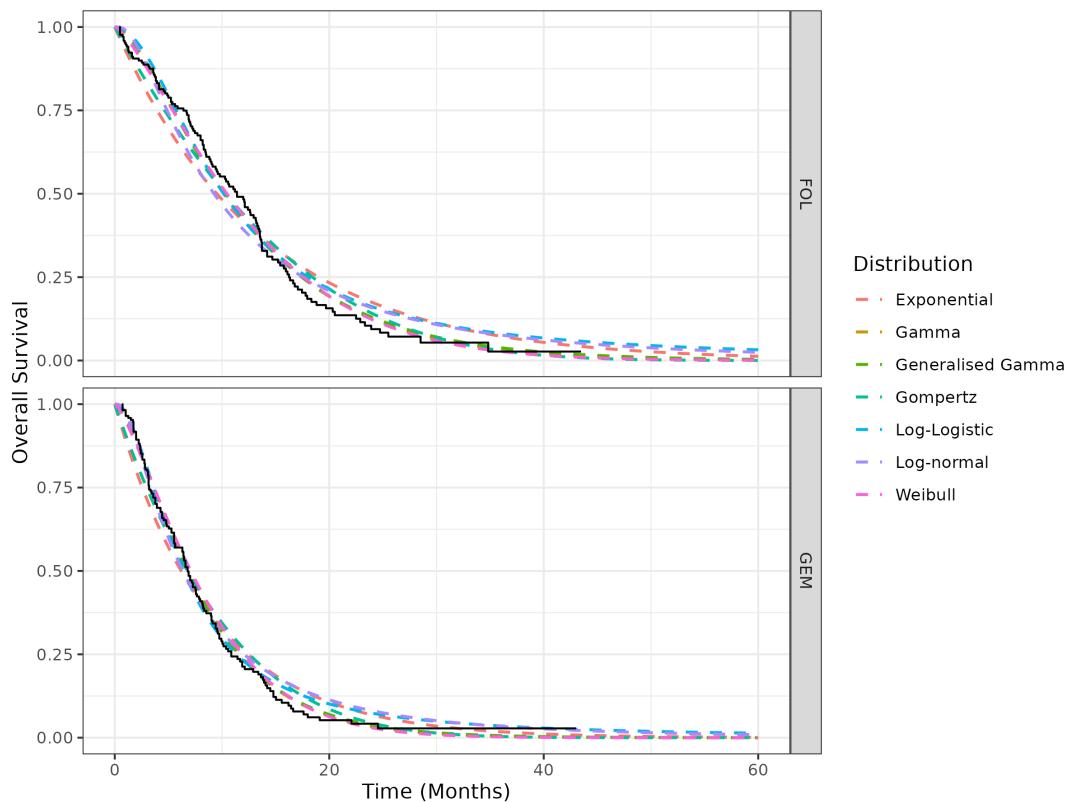


Figure 4.6: Conroy (2011) parametric model extrapolations

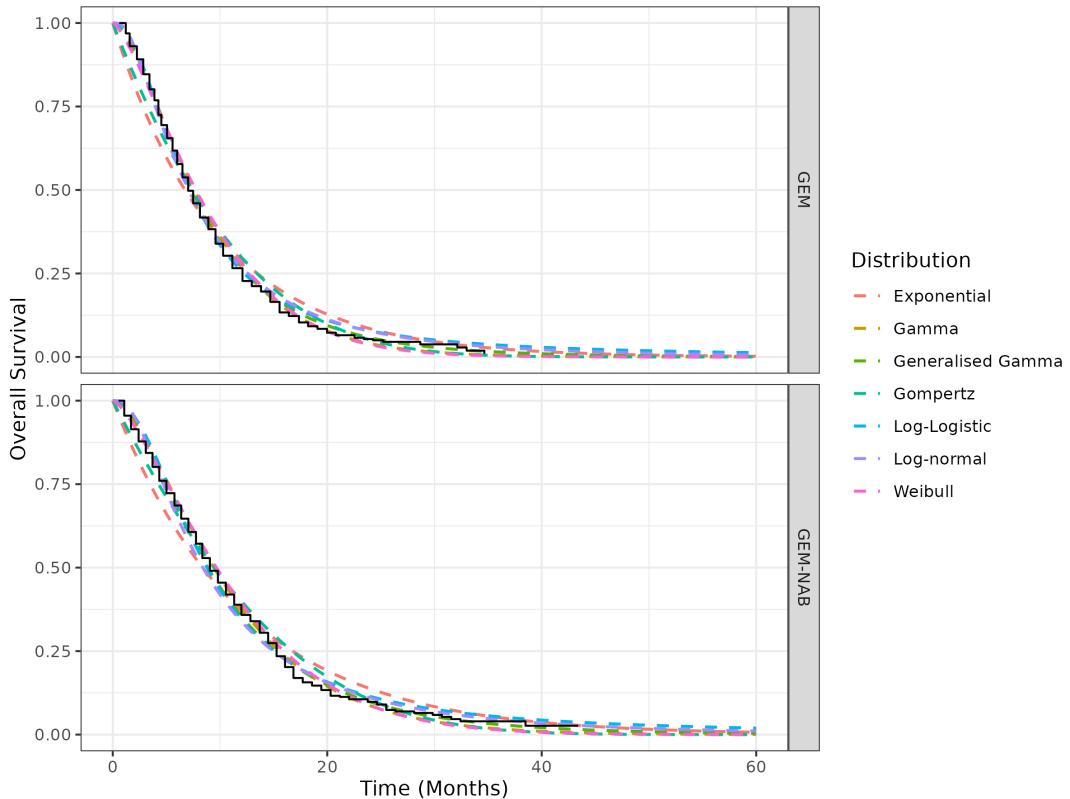


Figure 4.7: Goldstein (2015) parametric model extrapolations

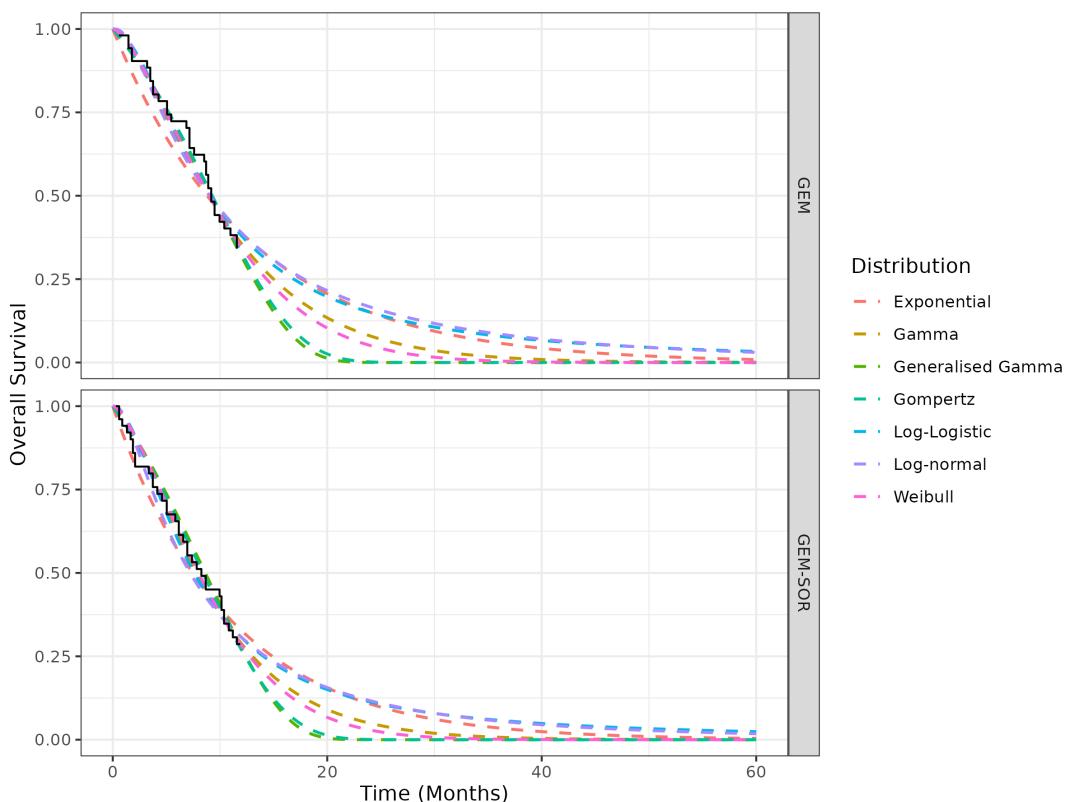


Figure 4.8: Goncalves (2012) parametric model extrapolations

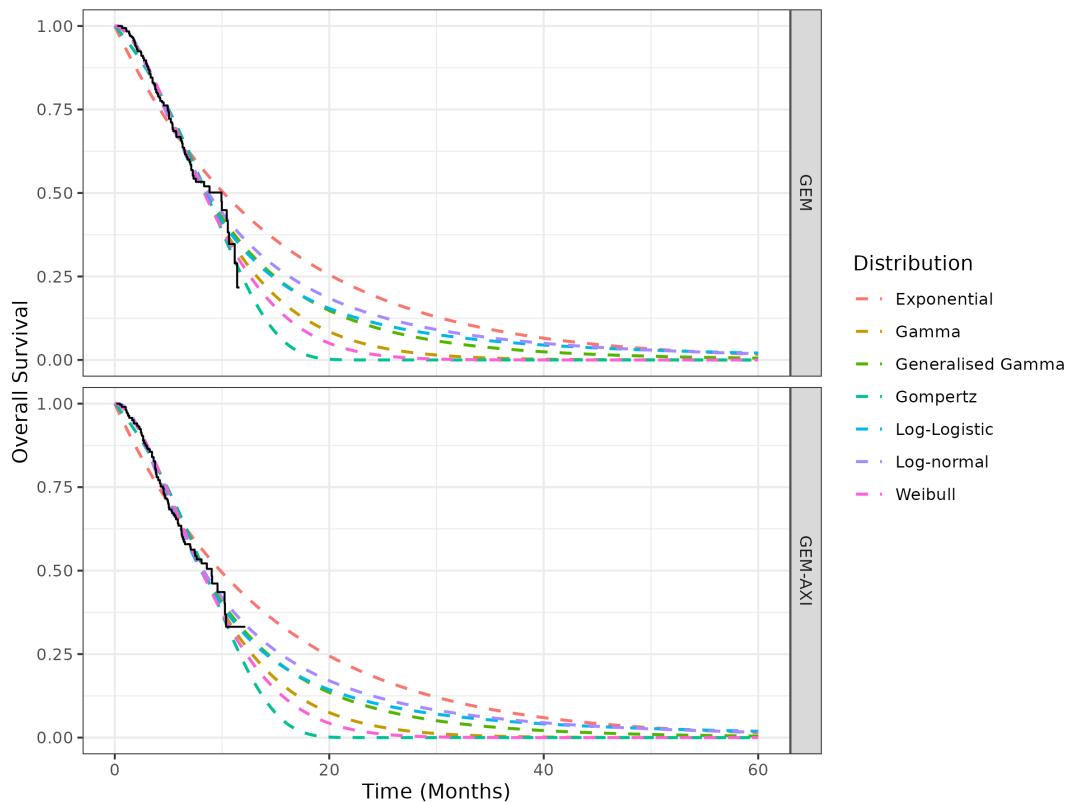


Figure 4.9: Kindler (2011) parametric model extrapolations

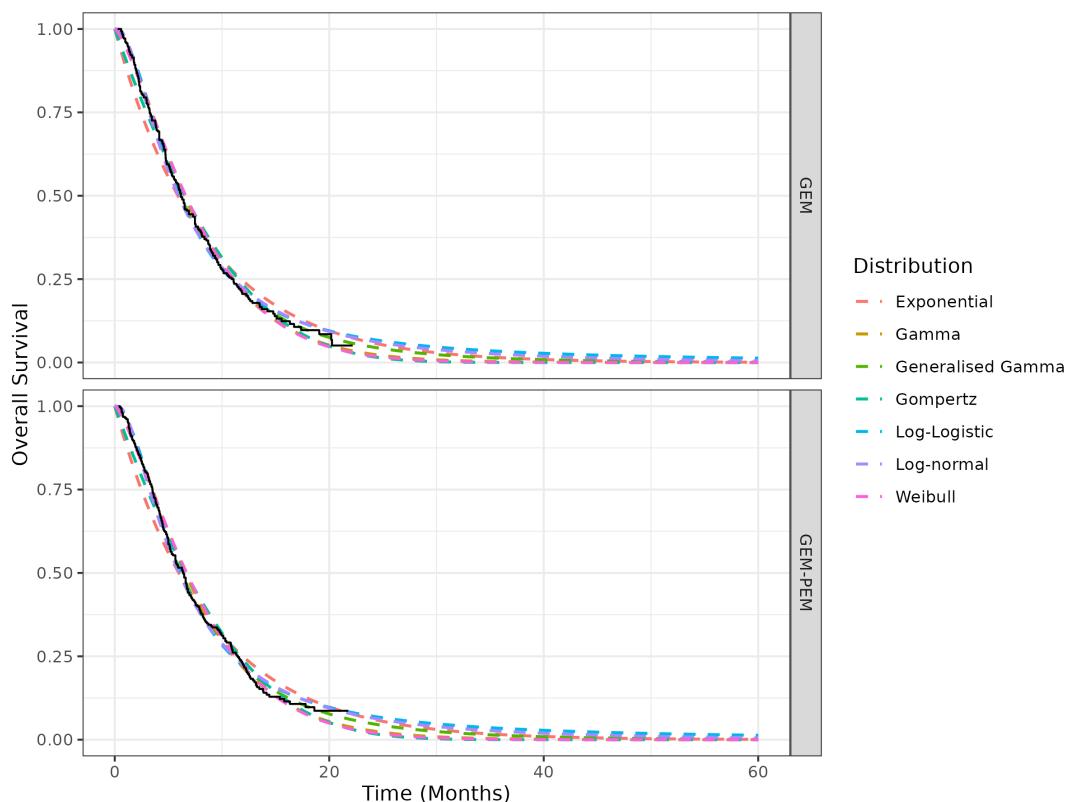


Figure 4.10: Oettle (2005) parametric model extrapolations

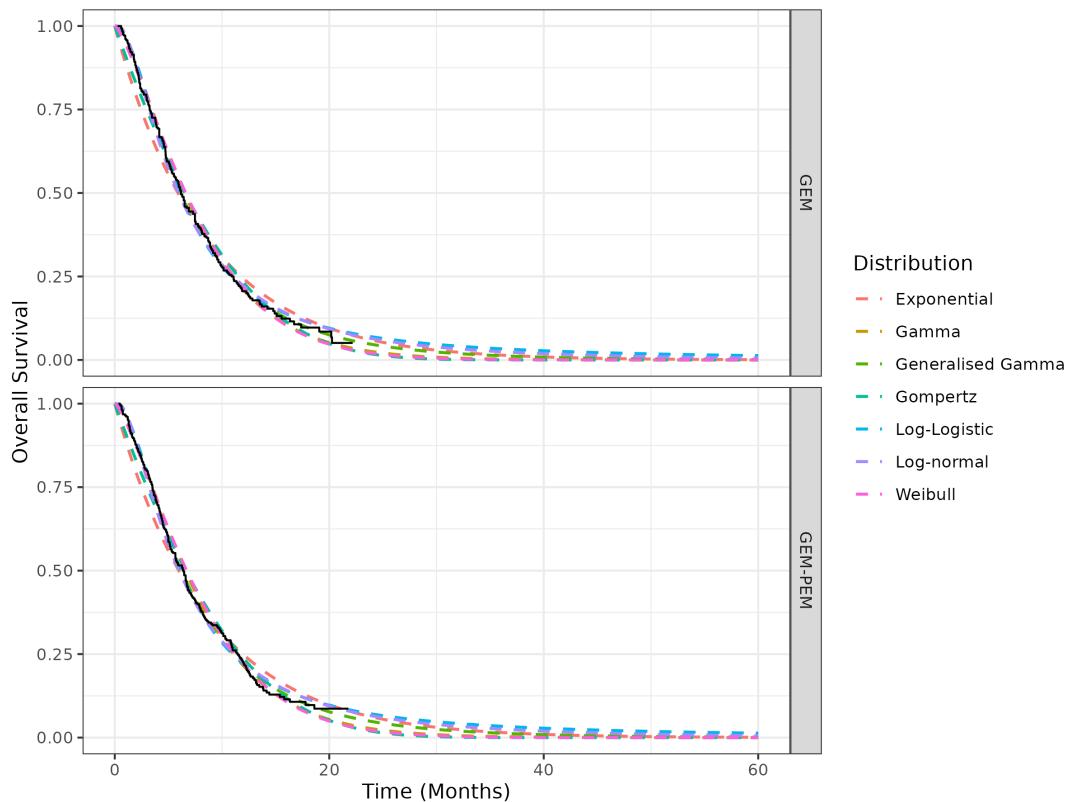


Figure 4.11: Rocha Lima (2004) parametric model extrapolations

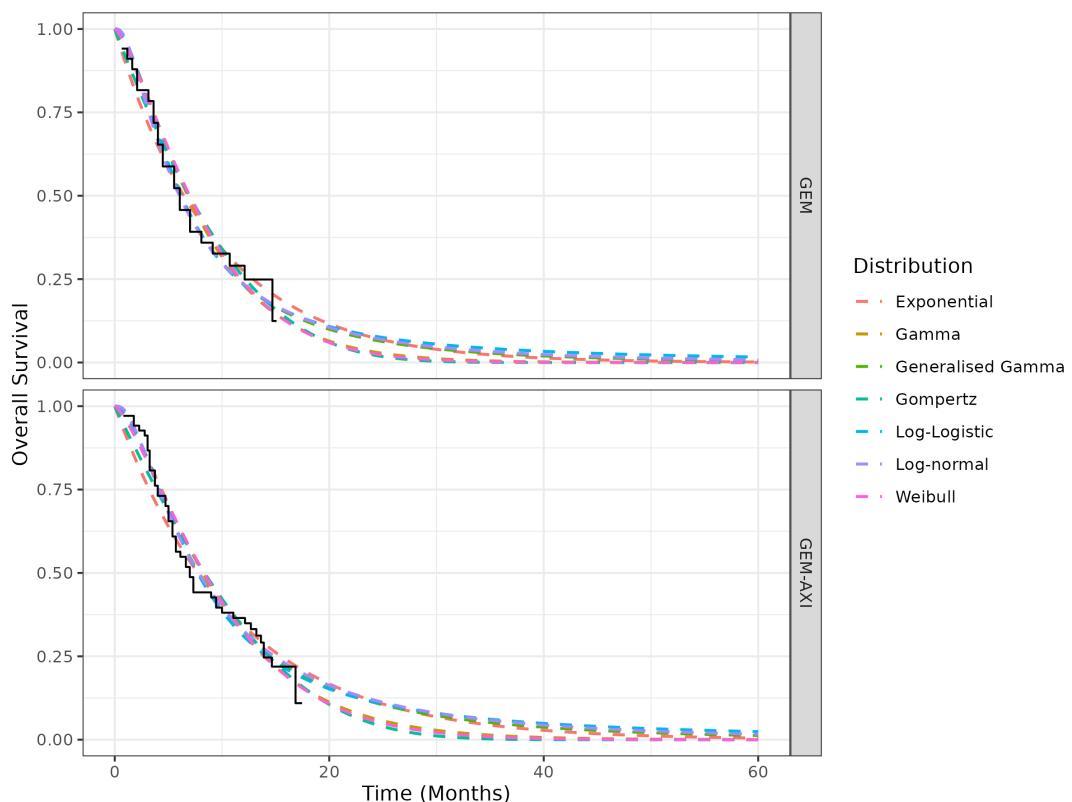


Figure 4.12: Spano (2008) parametric model extrapolations

A REVIEW OF NMAS FOR PANCREATIC CANCER

Prior to conducting this NMA, the literature was searched to identify NMAs comparing treatments for pancreatic cancer that included either all, or a subset of, the treatments included in this NMA.

In [Chen et al., 2021], the authors used a frequentist model to analyse both survival and toxicity data between modified FOLFIRINOX, regular FOLFIRINOX, and GEM-NAB. Data from 22 studies were included. The results of this NMA suggest that GEM-NAB and FOLFIRINOX are similarly efficacious, but FOLFIRINOX provides superior OS compared to GEM than GEM-NAB does. The results suggested that the improved efficacy of FOLFIRINOX and GEM-NAB were statistically significant, but GEM-NAB and FOLFIRINOX were not significantly different.

The NMA conducted in [Lin et al., 2019] included 31 studies in an efficacy meta-analysis and 32 studies in a safety meta-analysis. Naturally, only the efficacy analysis was relevant for this NMA. As with the Chen NMA, the authors used a frequentist NMA, and assessed the HRs. The focus of the Lin NMA was the first-line treatment of advanced pancreatic cancer. The only treatment included in the Lin NMA that was also relevant to this NMA was FOLFIRINOX. They found that FOLFIRINOX-based therapy was the best treatment.

The Nichetti meta-analysis [Nichetti et al., 2024] was not an NMA, but instead reconstructed KM data from seven studies, and then pooled the reconstructed IPD to compare the efficacy of FOLFIRINOX, NALIRIFOX, and GEM-NAB. They found that FOLFIRINOX provided superior OS to GEM-NAB. They did validate their results with an NMA, although this was not reported in the main paper. The median OS of all treatments in this study was below 12 months.

An NMA conducted by [Gresham et al., 2014], which included FOLFIRINOX, GEM-NAB, and GEM-CAP, found all three treatments to be associated with statistically significant improvements in OS relative to GEM and several other treatments. Their NMA was a Bayesian NMA for calculating survival outcomes. The primary result outcomes of their NMA were the HR and survival gain, as defined as in Equation 5.1.

$$\frac{\frac{\text{GEM Median OS}}{\text{HR}} - \text{GEM Median OS}}{\frac{\text{GEM Median PFS}}{\text{HR}} - \text{GEM Median PFS}} \quad (5.1)$$

The Takumoto [Takumoto et al., 2022] study considered only first line patients in Japan. They found FOLFIRINOX and GEM-NAB offered improved OS compared to GEM. The reported HRs indicated that FOLFIRINOX was more efficacious than GEM-NAB. The Takumoto NMA was performed in a Bayesian framework, using an RE model.

The [Zhang et al., 2018] NMA included several treatments that were also included in this NMA. They found FOLFIRINOX was the best treatment in the network. They used a RE model and summarised the results using the Surface Under the Cumulative Ranking Score (SUCRA) values. In terms of SUCRA values of 12-month OS, FOLFIRINOX performed best, followed by GEM+S1 (not included in this NMA) and GEM-NAB.

NMA OF PANCREATIC CANCER TRIALS

6.1 Network of Evidence

Figure 6.1 presents the network of evidence for this NMA. There were two studies comparing GEM with GEM-AXI, but only one study for each other comparison. As indicated by the size of each node, GEM-SOR was the treatment with the lowest sample size, and GEM-NAB was the comparator with the highest. The GEM → GEM-NAB edge is a different colour due to being an IPD trial.

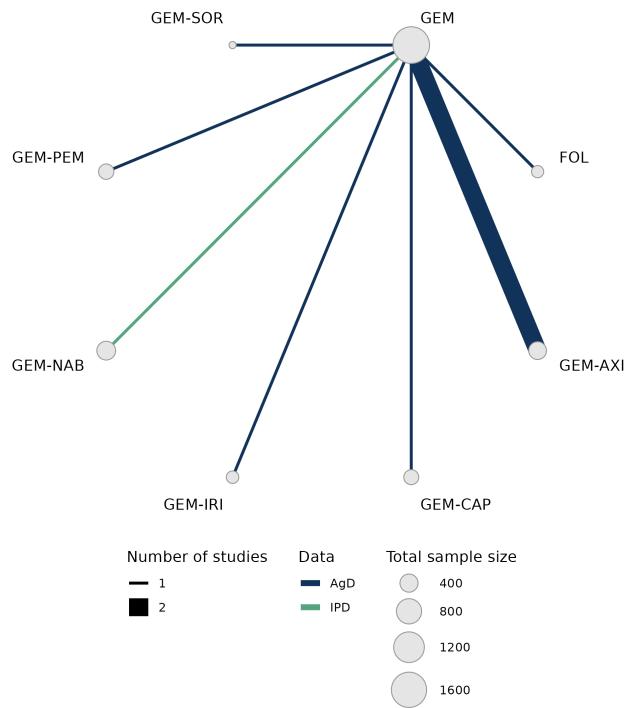


Figure 6.1: Network of evidence

6.2 Model Fitting and Selection

Both FE and RE models were fit using gamma, generalised gamma, Gompertz, log-logistic, log-normal, and Weibull likelihoods. Vague priors were used for each model. Namely, the intercept prior was $N(0, 100)$, the treatment prior was $N(0, 10)$, the auxiliary prior was $hN(0, 5)$, and auxiliary regression prior was $N(0, 10)$. Here, hN denotes a *half-normal distribution*, as defined in Definition ???. For each model, sampling was done using 1000 iterations on four chains. The first 500 iterations were warmup

iterations.

Table 6.1 presents the selection statistics for each model. The fixed effect model was also deemed to be clinically appropriate due to homogeneity in the patient population. The FE gamma, FE and RE generalised gamma models failed to converge, meaning no DIC or LOOIC estimates could be obtained. The RE gamma and RE Gompertz models gave LOOIC and DIC scores so high that they were classified as Inf, indicating poor fit. The log-logistic, log-normal, and Weibull models gave similar LOOIC and DIC scores. In each case, the FE model had slightly lower LOOIC than the RE model for each likelihood. The FE log-normal model gave the lowest LOOIC and DIC score, indicating it was the best fitting model. The trace plot for this model is available in Appendix A, Figure A.1, and indicated good convergence due to consistent peaks and troughs across each treatment arm.

Likelihood	Effect	DIC	LOOIC
Log-logistic	Fixed	15176.9789	15176.0719
Log-logistic	Random	15175.2325	15176.2254
Log-normal	Fixed	15172.3036 ←	15173.0542 ←
Log-normal	Random	15172.7801	15173.3975
Weibull	Fixed	15200.20470	15200.1331
Weibull	Random	15199.29711	15200.6538

Table 6.1: Model selection statistics for each model

6.3 Results

Figure ?? and Figure ?? present the predicted survival and hazard of each treatment in study population, respectively. GEM-NAB and GEM-CAP had the highest OS and lowest hazard in each study population. The hazard curves for each study in each population followed a similar pattern, with peaks in the hazard just before ten months, before declining. GEM-SOR had the highest peak-hazard in each population, but crossed the GEM hazard curve in each population shortly after the peak, finishing with a lower hazard than GEM by the end of the observation period in each population. Further, in terms of hazard, the GEM-AXI and GEM-PEM curves were almost identical in each population.

Figure 6.4 presents the estimated RMST of each treatment in each population. The GEM-SOR arm had large credible intervals in each population due to the lower number of patients for which GEM-SOR data was available. Further, GEM-SOR and GEM-IRI had the lowest and second-lowest RMST estimates in each population, respectively. The RMST estimates for GEM, GEM-AXI, and GEM-PEM were similar in each population, as were GEM-CAP, and GEM-NAB. GEM-NAB and GEM-CAP had the highest and second-highest RMST estimates respectively in every study population.

Figure 6.5 presents the estimated median OS of each treatment in each population. The median OS estimates follow the same pattern as the RMST estimates. Namely, GEM-SOR and GEM-IRI gave the lowest and second-lowest estimates for median OS in each study population, GEM, GEM-AXI, and GEM-PEM gave similar estimates, and GEM-NAB and GEM-CAP gave the highest and second-highest estimates of median OS, respectively. The median OS estimates of GEM-NAB and GEM-CAP were further apart than the RMST estimates for the same two treatments.

Figure 6.2: OS of each treatment in each population

Figure 6.3: Hazards of each treatment in each population

Figure 6.4: RMST of each treatment in each population

Figure 6.5: Median OS of each treatment in each population

CONCLUSION AND DISCUSSIONS

7.1 Conclusion

7.2 Discussion and Further Work

The results obtained in this dissertation align with an NMA conducted by [Gresham et al., 2014], which also found GEM-NAB and GEM-CAP offered improvements in OS compared to GEM monotherapy. This NMA included fewer studies than that of [Gresham et al., 2014], so future work to expand the number of trials would be good to solidify the place of GEM-NAB and GEM-CAP as the best treatment options for advanced/pancreatic cancer.

In 2017, NICE published TA476 [National Institute for Health and Care Excellence, 2017], which recommended GEM-NAB as an option for untreated metastatic pancreatic cancer in adults only if other combination chemotherapies are unsuitable and the patient would otherwise receive GEM monotherapy. The TA claimed that GEM-NAB was more effective in increasing than GEM monotherapy, similarly effective to GEM-CAP, but was less effective than FOLFIRINOX. The comparison to GEM-CAP was stated as “uncertain”.

Given the poor prognosis of pancreatic cancer, even a small improvement in median OS of a couple of months is of immense emotional value to patients and their families, and should not be overlooked for looking like small improvements out of context.

7.3 Considerations for the ISPOR Good Practice Task Force

ISPOR developed a 26-item questionnaire for assessing the credibility of an NMA [Jansen et al., 2014]. This NMA was performed inline with these practices. While each question is not answered individually here, the themes of the guidance, and how this NMA aligns with it, are discussed. An answer to each question individually is available in Appendix B

The first set of questions in the guidance concerns the evidence base. This NMA was performed on a fully-connected network of evidence (Figure 6.1), and included no poor-quality studies. Indeed, the study populations and trial characteristics were similar, meaning there was no systematic differences in treatment effect modifiers across the comparisons. The only aspect of this NMA that could be considered not to follow these guidelines was that not all available RCTs were included. The Greshem study, for example included 23 studies obtained by searching several databases. This NMA was not conducted based on results of a systematic literature review or database search. Studies were selected for this NMA based on a brief literature search for trials comparing GEM with another therapy. Since all the KM curves from published papers had to be digitised, which takes a considerable amount of time, there was always to be a limit on how many studies could be included. Since pancreatic cancer treatments have not changed much, the fact the studies included in this NMA were reported between 2004 and 2015 was not deemed to render the results of the NMA inconclusive in today’s treatment landscape.

The second set of questions concerns the analysis. No naïve comparisons were made, which preserve within-study randomisation. As there were no cases of both direct and indirect evidence for any

treatments, questions eight and nine were not deemed relevant. Question ten concerns imbalance of the distribution of effect modifiers, and how this was accounted for. Since the ML-NMR is a meta-regression model, this was directly accounted for. In terms of FE and RE models, both were fit, and the best fitting model selected in terms of robust selection statistics. Since the studies included in this NMA were not diverse in terms of methodology, FE models were deemed to be clinically appropriate. The guidance generally recommends RE models, but it was deemed clinically appropriate to consider FE models in this case. Were more trials to be included, more consideration would need to be given to the similarity assessment to determine the suitability of FE models.

The third set of questions relates to the reporting quality. While all the studies used, and indeed the associated KM curves were presented, the actual TTE data was not presented. This is due to the form of the data, although it is available within PCNMA R package. Individual study results were provided in Figures 6.2- 6.5. Considerations did not have to be made for direct and indirect comparisons since there were no closed loops. Rankings were reported to address the main project aim, and pairwise comparisons were reported. In particular, the pairwise comparison between GEM-CAP and GEM-NAB was important to clarify the uncertainty mentioned by NICE in NG85. No consideration was given to the effect of important patient characteristics due to the homogeneity in the trial populations and further because of a lack of IPD available for this study. It is not sound for those involved in the study to assess the fairness of the conclusions and interpretation, but every attempt was made to perform this NMA with integrity and interpret the results in line with the evidence.

BIBLIOGRAPHY

- [Blackford et al., 2009] Blackford, A., Parmigiani, G., Kensler, T. W., Wolfgang, C., Jones, S., Zhang, X., Parsons, D. W., Lin, J. C.-H., Leary, R. J., Eshleman, J. R., et al. (2009). Genetic mutations associated with cigarette smoking in pancreatic cancer. *Cancer research*, 69(8):3681–3688.
- [Cancer Research UK, 2024] Cancer Research UK (2024). Pancreatic cancer statistics.
- [Chen et al., 2021] Chen, J., Hua, Q., Wang, H., Zhang, D., Zhao, L., Yu, D., Pi, G., Zhang, T., and Lin, Z. (2021). Meta-analysis and indirect treatment comparison of modified folfirinox and gemcitabine plus nab-paclitaxel as first-line chemotherapy in advanced pancreatic cancer. *BMC cancer*, 21:1–9.
- [Conroy et al., 2011] Conroy, T., Desseigne, F., Ychou, M., Bouché, O., Guimbaud, R., Bécouarn, Y., Adenis, A., Raoul, J.-L., Gourgou-Bourgade, S., De La Fouchardière, C., et al. (2011). Folfirinox versus gemcitabine for metastatic pancreatic cancer. *New England journal of medicine*, 364(19):1817–1825.
- [Cox, 1972] Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2):187–202.
- [Cunningham et al., 2009] Cunningham, D., Chau, I., Stocken, D. D., Valle, J. W., Smith, D., Steward, W., Harper, P. G., Dunn, J., Tudur-Smith, C., West, J., et al. (2009). Phase iii randomized comparison of gemcitabine versus gemcitabine plus capecitabine in patients with advanced pancreatic cancer. *Journal of Clinical Oncology*, 27(33):5513–5518.
- [Deeb et al., 2015] Deeb, A., Haque, S.-U., and Olowokure, O. (2015). Pulmonary metastases in pancreatic cancer, is there a survival influence? *Journal of gastrointestinal oncology*, 6(3):E48.
- [Goldstein et al., 2015] Goldstein, D., El-Maraghi, R. H., Hammel, P., Heinemann, V., Kunzmann, V., Sastre, J., Scheithauer, W., Siena, S., Tabernero, J., Teixeira, L., et al. (2015). nab-paclitaxel plus gemcitabine for metastatic pancreatic cancer: long-term survival from a phase iii trial. *Journal of the National Cancer Institute*, 107(2):dju413.
- [Gonçalves et al., 2012] Gonçalves, A., Gilabert, M., François, E., Dahan, L., Perrier, H., Lamy, R., Re, D., Largillier, R., Gasmi, M., Tchiknavorian, X., et al. (2012). Baypan study: a double-blind phase iii randomized trial comparing gemcitabine plus sorafenib and gemcitabine plus placebo in patients with advanced pancreatic cancer. *Annals of oncology*, 23(11):2799–2805.
- [Gresham et al., 2014] Gresham, G. K., Wells, G. A., Gill, S., Cameron, C., and Jonker, D. J. (2014). Chemotherapy regimens for advanced pancreatic cancer: a systematic review and network meta-analysis. *BMC cancer*, 14:1–13.
- [Guyot et al., 2012] Guyot, P., Ades, A., Ouwens, M. J., and Welton, N. J. (2012). Enhanced secondary analysis of survival data: reconstructing the data from published kaplan-meier survival curves. *BMC medical research methodology*, 12:1–13.
- [Jackson, 2016] Jackson, C. H. (2016). flexsurv: a platform for parametric survival modeling in r. *Journal of statistical software*, 70.
- [Jansen et al., 2014] Jansen, J. P., Trikalinos, T., Cappelleri, J. C., Daw, J., Andes, S., Eldessouki, R., and Salanti, G. (2014). Indirect treatment comparison/network meta-analysis study questionnaire to assess relevance and credibility to inform health care decision making: an ispor-amcp-npc good practice task force report. *Value in Health*, 17(2):157–173.

- [Kelsen et al., 1997] Kelsen, D. P., Portenoy, R., Thaler, H., Tao, Y., and Brennan, M. (1997). Pain as a predictor of outcome in patients with operable pancreatic carcinoma. *Surgery*, 122(1):53–59.
- [Kindler et al., 2011] Kindler, H. L., Ioka, T., Richel, D. J., Bennouna, J., Létourneau, R., Okusaka, T., Funakoshi, A., Furuse, J., Park, Y. S., Ohkawa, S., et al. (2011). Axitinib plus gemcitabine versus placebo plus gemcitabine in patients with advanced pancreatic adenocarcinoma: a double-blind randomised phase 3 study. *The lancet oncology*, 12(3):256–262.
- [Latimer, 2011] Latimer, N. (2011). Nice dsu technical support document 14: survival analysis for economic evaluations alongside clinical trials-extrapolation with patient-level data. *Report by the Decision Support Unit*.
- [Lin et al., 2019] Lin, K.-I., Yang, J.-L., Lin, Y.-C., Chou, C.-Y., Chen, J.-H., and Hung, C.-C. (2019). Network meta-analysis of efficacy and safety of chemotherapy and target therapy in the first-line setting of advanced pancreatic cancer. *Cancers*, 11(11):1746.
- [National Institute for Health and Care Excellence, 2017] National Institute for Health and Care Excellence (2017). Paclitaxel as albumin-bound nanoparticles with gemcitabine for untreated metastatic pancreatic cancer.
- [National Institute for Health and Care Excellence, 2018] National Institute for Health and Care Excellence (2018). Pancreatic cancer in adults: diagnosis and management.
- [Nichetti et al., 2024] Nichetti, F., Rota, S., Ambrosini, P., Pircher, C., Gusmaroli, E., Busset, M. D. D., Pusceddu, S., Sposito, C., Coppa, J., Morano, F., et al. (2024). Nalirifox, folirinox, and gemcitabine with nab-paclitaxel as first-line chemotherapy for metastatic pancreatic cancer: a systematic review and meta-analysis. *JAMA Network Open*, 7(1):e2350756–e2350756.
- [Oettle et al., 2005] Oettle, H., Richards, D., Ramanathan, R., Van Laethem, J.-L., Peeters, M., Fuchs, M., Zimmermann, A., John, W., Von Hoff, D., Arning, M., et al. (2005). A phase iii trial of pemetrexed plus gemcitabine versus gemcitabine in patients with unresectable or metastatic pancreatic cancer. *Annals of Oncology*, 16(10):1639–1645.
- [Pancreatic Cancer UK, 2022] Pancreatic Cancer UK (2022). Signs and symptoms of pancreatic cancer.
- [Phillippo et al., 2020] Phillippe, D. M., Dias, S., Ades, A., Belger, M., Brnabic, A., Schacht, A., Saure, D., Kadziola, Z., and Welton, N. J. (2020). Multilevel network meta-regression for population-adjusted treatment comparisons. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 183(3):1189–1210.
- [Phillippo et al., 2024] Phillippe, D. M., Dias, S., Welton, N. J., and Ades, A. (2024). Multilevel network meta-regression for general likelihoods: synthesis of individual and aggregate data with applications to survival analysis. *arXiv preprint arXiv:2401.12640*.
- [Rocha Lima et al., 2004] Rocha Lima, C. M., Green, M. R., Rotche, R., Miller Jr, W. H., Jeffrey, G. M., Cisar, L. A., Morganti, A., Orlando, N., Gruia, G., and Miller, L. L. (2004). Irinotecan plus gemcitabine results in no survival advantage compared with gemcitabine monotherapy in patients with locally advanced or metastatic pancreatic cancer despite increased tumor response rate. *Journal of Clinical Oncology*, 22(18):3776–3783.
- [Sobol', 1967] Sobol', I. M. (1967). On the distribution of points in a cube and the approximate evaluation of integrals. *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki*, 7(4):784–802.
- [Spano et al., 2008] Spano, J.-P., Chodkiewicz, C., Maurel, J., Wong, R., Wasan, H., Barone, C., Létourneau, R., Bajetta, E., Pithavala, Y., Bycott, P., et al. (2008). Efficacy of gemcitabine plus axitinib compared with gemcitabine alone in patients with advanced pancreatic cancer: an open-label randomised phase ii study. *The Lancet*, 371(9630):2101–2108.
- [Takumoto et al., 2022] Takumoto, Y., Sasahara, Y., Narimatsu, H., and Akazawa, M. (2022). Comparative outcomes of first-line chemotherapy for metastatic pancreatic cancer among the regimens used in japan: a systematic review and network meta-analysis. *JAMA Network Open*, 5(1):e2145515–e2145515.
- [Vehtari et al., 2024] Vehtari, A., Gabry, J., Magnusson, M., Yao, Y., Bürkner, P.-C., Paananen, T., and Gelman, A. (2024). loo: Efficient leave-one-out cross-validation and waic for bayesian models. R package version 2.7.0.

- [Vehtari et al., 2017] Vehtari, A., Gelman, A., and Gabry, J. (2017). Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and computing*, 27:1413–1432.
- [Wickham and Henry, 2023] Wickham, H. and Henry, L. (2023). *purrr: Functional Programming Tools*. R package version 1.0.2, <https://github.com/tidyverse/purrr>.
- [Wong et al., 2009] Wong, A., Soo, R. A., Yong, W.-P., and Innocenti, F. (2009). Clinical pharmacology and pharmacogenetics of gemcitabine. *Drug metabolism reviews*, 41(2):77–88.
- [Zhang et al., 2018] Zhang, S.-H., Liu, G.-F., Li, X.-F., Liu, L., and Yu, S.-N. (2018). Efficacy of different chemotherapy regimens in treatment of advanced or metastatic pancreatic cancer: a network meta-analysis. *Journal of Cellular Physiology*, 233(4):3352–3374.

Appendices

ADDITIONAL NMA RESULTS

This appendix presents some additional NMA results for the FE log-normal ML-NMR model.

A.1 Model Convergence

Figure A.1: Trace plot for the FE log-normal model ML-NMR

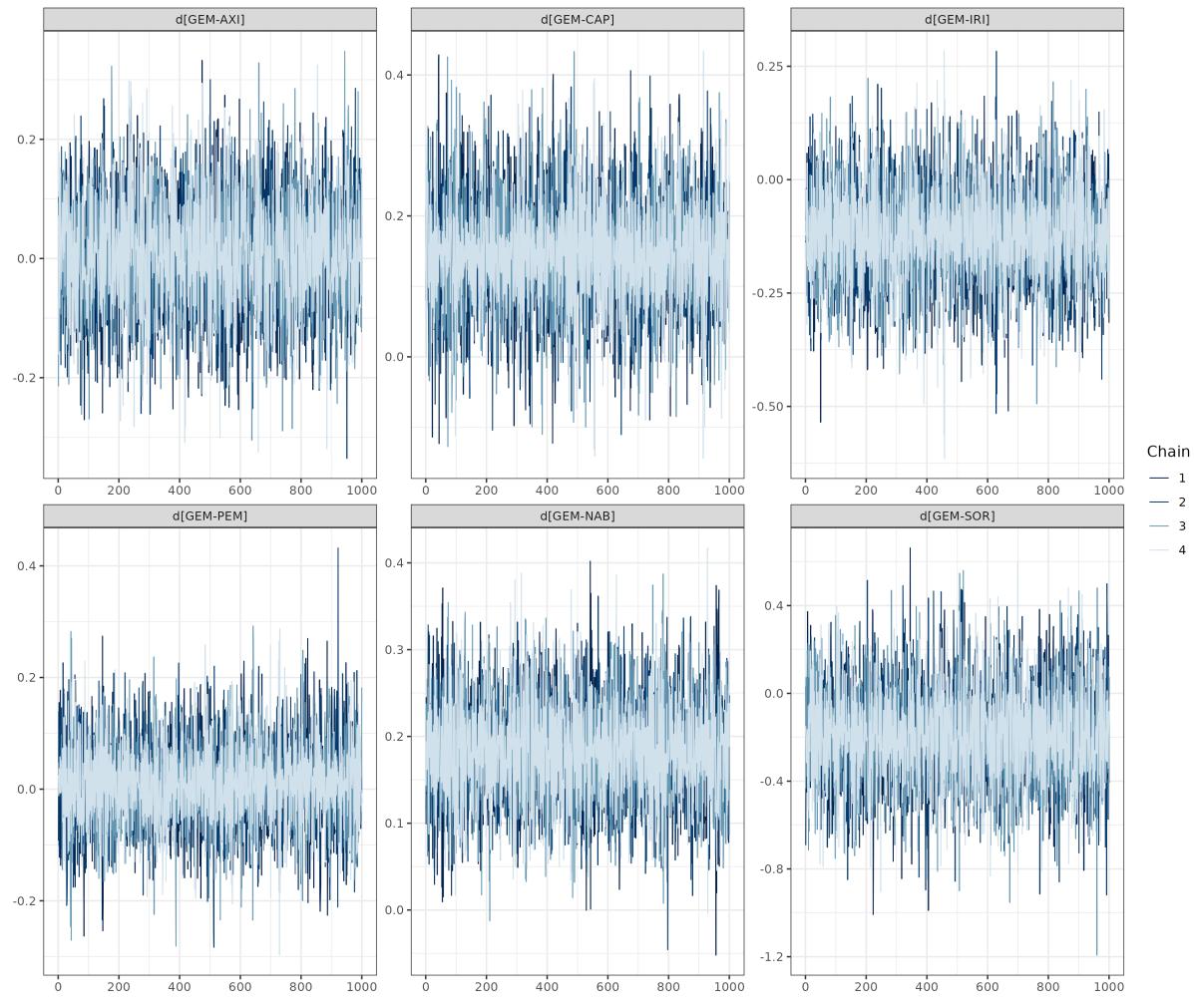


Figure A.2: Trace plot for the FE log-normal model ML-NMR

ISPOR GOOD PRACTICE QUESTIONS

B.1 Evidence Base

Is the population relevant? Yes. The populations in these trials were reflective of people most likely to have advanced/metastatic pancreatic cancer.

Are any relevant interventions missing? Potentially. The Gresham NMA included 19 treatments in total. This NMA included all of the best-performing treatments from the Gresham study, but could be expanded to include more studies in the future. The treatments in this NMA however were chosen as the ones most likely to be of interest to a decision maker in a clinical setting.

Are any relevant outcomes missing? No PFS data was included. The purpose of this NMA was for assessing the OS endpoint, but the same methods would apply with PFS data.

Is the context (settings and circumstances) applicable? Treatment for pancreatic cancer has not changed much over the last 25 years. Indeed, while the studies used in this NMA were reported between 2004 and 2015, all the treatments are still relevant in 2024. If surgical studies had been included, there may be some discussion required, as more attempts are being made to operate on pancreatic cancer nowadays.

Did the researchers attempt to identify and include all relevant RCTs? No. no thorough literature search was performed for this NMA. Studies were selected based on how well they were reported, due to considerations with digitising the KM curves, and whether the comparators were relevant. As mentioned previously, the best-performing treatments from the Gresham study were of primary interest for this NMA.

Do the trials for the interventions of interest form one connected network of RCTs? Yes, see Figure 6.1.

Is it apparent that poor quality studies were included, thereby leading to bias? No. The only potential comment here is the low sample size of the Spano or Goncalves studies. These studies were not deemed to be of poor quality despite this.

Is it likely that bias was induced by selective reporting of outcomes in the studies? As only the OS endpoint was considered, this was not deemed to be a potential influencer of any bias in the NMA.

Are there systematic differences in treatment effect modifiers (i.e., baseline patient or study characteristics that have an impact on the treatment effects) across the different treatment comparisons in the network? No. The populations in the included studies were more-or-less identical. We included the proportion of male patients as a covariate in order to use the ML-NMR method.

If yes (i.e., there are such systematic differences in treatment effect modifiers), were these imbalances in effect modifiers across the different treatment comparisons identified

before comparing individual study results? NA.

B.2 Analysis

Were statistical methods used that preserve within-study randomization? (No naive comparisons) Yes. The NMA was based on relative treatment effects.

If both direct and indirect comparisons are available for pairwise contrasts (i.e., closed loops), was agreement in treatment effects (i.e., consistency) evaluated or discussed? No closed loops. Question not applicable.

In the presence of consistency between direct and indirect comparisons, were both direct and indirect evidence included in the network meta-analysis? No closed loops. Question not applicable.

With inconsistency or an imbalance in the distribution of treatment effect modifiers across the different types of comparisons in the network of trials, did the researchers attempt to minimize this bias with the analysis? This was not deemed relevant due to the similarity of the trials.

Was a valid rationale provided for the use of random-effects or fixed-effect models? Given the similarity of the included trials, FE models were not deemed to be clinically inappropriate. This is why both FE and RE models were fit for each likelihood. The best performing models were selected based on the LOOIC and DIC scores, rather than any clinical considerations.

If a random-effects model was used, were assumptions about heterogeneity explored or discussed? MATT: ANSWER THIS AFTER WRITING RESULTS!!

If there are indications of heterogeneity, were subgroup analyses or meta-regression analysis with prespecified covariates performed? MATT: ANSWER THIS AFTER WRITING RESULTS!!

B.3 Reporting quality and transparency

Is a graphical or tabular representation of the evidence network provided with information on the number of RCTs per direct comparison? Yes, see Figure 6.1. The thickness of lines denotes the number of RCTs available per comparison.

Are the individual study results reported?

Are results of direct comparisons reported separately from results of the indirect comparisons or network meta-analysis?

Are all pairwise contrasts between interventions as obtained with the network meta-analysis reported along with measures of uncertainty?

Is a ranking of interventions provided given the reported treatment effects and its uncertainty by outcome?

Is the effect of important patient characteristics on treatment effects reported?

B.4 Interpretation

Are the conclusions fair and balanced? Yes. Every attempt was made to do this.

B.5 Conflicts of Interest

Were there any potential conflicts of interest? This project was somewhat personal to the author, but that personal experience did not influence the results. There was no inherent bias to a

particular therapy. The author is employed by a health economics consultancy, but has no interest in a particular therapy from a commercial perspective either.

If yes, were steps taken to address these? Not required.

THE PCNMA PACKAGE

To facilitate the analysis conducted in this project, an R package was constructed. Performing the analysis in this way ensures easy reproducibility for further analysis in the future. The package is roughly split into two parts: survival code, and NMA code.

C.1 Survival Functions

The central function in the package for survival analysis is the *fit_distribution* function. This function is a wrapper around the *flexsurvreg* function from *flexsurv*. It is designed to take a list object containing the names of distributions and the associated argument to pass that distribution to *flexsurv*. For example,

```
1 distributions <- list("Weibull" = "weibull")
```

The reason for defining the distributions in this way is because the default *flexsurv* arguments look untidy by default when plotting. For example, the exponential distribution is passed as just “exp”. The beauty of the *fit_distribution* function is that several distributions can be passed with a tidier name, which *flexsurvreg* never sees. This purely aesthetic change can be seen by comparing Figure C.1, and Figure C.2.

Any default *flexsurv* distribution will work with *fit_distribution*, it just needs a suitable name in the list object . The *fit_distribution* function itself doesn’t actually fit any distributions. There is a sub-function, *.fit_distribution* which takes a single distribution as an argument, and some data on which to fit that distribution. This function takes the “weibull” element of the above list and passes it to *flexsurv*. The *fit_distribution* function maps *.fit_distribution* across the list of distributions, using the *purrrr::map* function [Wickham and Henry, 2023]. After some data cleaning, the object that is returned by *fit_distribution* is given the class “*fitted_distribution*”.

Several S3 methods exist for objects of class “*fitted_distribution*”, these are *plot.fitted_distribution*, *summary.fitted_distribution*, and *coef.fitted_distribution*. These functions allow for plotting fitted models, accessing information such as AIC scores, and accessing the model coefficients respectively.

C.2 NMA Functions

The NMA part of the PCNMA package is a wrapper around *multinma*. In particluar, the *fit_model* function is a wrapper around *multinma::nma*.

Package ‘PCNMA’

June 16, 2024

Type Package

Title Parametric and Non-Parametric Survival Models in Pancreatic Cancer Trials

Version 0.1

Date 2023-06-27

Author Matthew Knowles

Maintainer Matthew Knowles <mattknowles314@gmail.com>

Description This package contains the functions and data for my MSc Thesis, concerning the use of parametric and non-parametric models for survival in pancreatic cancer

License MIT + file LICENSE

RoxygenNote 7.2.3

Encoding UTF-8

R topics documented:

.fit_distribution	2
boxTid	2
coef.fitted_distribution	3
fit_distribution	3
fit_model	4
gen_network	4
gen_network_data	5
H	5
hr	6
km_estimates	6
phi	7
plot.fitted_distribution	7
plot.fitted_model	8
plot.km_obj	8
plot_network	9
rmst	10
summary.fitted_distribution	10
summary.fitted_model	11
summary.hr_obj	11
summary.km_obj	11

Index**12**

<code>.fit_distribution</code>	<i>Fit a single survival distribution</i>
--------------------------------	---

Description

This function is a wrapper on ‘flexsurv::flexsurvreg‘, and fits a single distribution to the specified data.

Usage

```
.fit_distribution(distribution, data, strata = "Treatment")
```

Arguments

<code>distribution</code>	A single distribution
<code>data</code>	An IPD dataset
<code>strata</code>	Strata for the RHS of the ‘survival::Surv‘ function

Value

A [flexsurv::flexsurvreg] object

<code>boxTid</code>	<i>Box-Tidwell Transformation</i>
---------------------	-----------------------------------

Description

Box-Tidwell Transformation

Usage

```
boxTid(x, p)
```

Arguments

<code>x</code>	A real value
<code>p</code>	The p-value to raise x to

```
coef.fitted_distribution
  Coefficients of fitted models
```

Description

Returns the model coefficients for a given survival model

Usage

```
## S3 method for class 'fitted_distribution'
coef(fit, studies, ...)
```

Arguments

fit	A [PCNMA::fitted_distribution] object
...	for S3 consistency

fit_distribution	<i>Fit survival distributions to a dataset.</i>
------------------	---

Description

This function extends the ‘PCNMA:::fit_distribution’ function, by fitting a given set of distributions to a TTE dataset.

Usage

```
fit_distribution(
  distributions = nice_parametric_dists,
  data,
  strata = "Treatment",
  maxT = 60
)
```

Arguments

distributions	A list of distributions
data	An IPD dataset
strata	Stratification variables
maxT	maximum time to calculate fitted values at

Value

A dataframe with fitted values

fit_model	<i>Run an NMA</i>
------------------	-------------------

Description

Run an NMA

Usage

```
fit_model(network, effects, seed = 1, chains = 4, llhood = "weibull")
```

Arguments

network	A [multinma::nma_data] object
llhood	Character string specifying a likelihood function
link	Character string specifying a link function (defaults to "log")
...	Other parameters to pass to [multinma::nma]

gen_network	<i>Generate a network of evidence</i>
--------------------	---------------------------------------

Description

Generate a network of evidence

Usage

```
gen_network(net_data, ref, covs)
```

Arguments

net_data	A dataset created by [PCNMA::gen_network_data]
ref	A character reference treatment

Value

A [mutlinma::nma_data] object

gen_network_data *Generate network data*

Description

Generate network data

Usage

gen_network_data(data, ref)

Arguments

data	A date extraction dataset
ref	A character reference treatment

Value

A dataframe

H *H-function for FPs*

Description

H-function for FPs

Usage

H(x, P, zeta, j)

Arguments

x	A real value
P	A vector of powers
zeta	A vector of zeta values
j	The index

hr	<i>Hazard Ratio</i>
----	---------------------

Description

Hazard Ratio

Usage

```
hr(TTE, strata = "1")
```

Arguments

TTE	A TTE dataframe
strata	A strata variable

km_estimates	<i>Generate KM estimates</i>
--------------	------------------------------

Description

Generate Kaplan-Meier estimates from a TTE object. Uses ‘ggsurvfit::survfit2‘ to generate KM estimates. Only requires strata to be specified.

Usage

```
km_estimates(TTE, strata = "1")
```

Arguments

TTE	A TTE dataframe
-----	-----------------

Value

A [PCNMA::km_obj] object

phi *Fractional Polynomial Function*

Description

Fractional Polynomial Function

Usage

```
phi(x, m, P, zeta)
```

Arguments

x	A real value
m	The degree of the polynomial
P	A vector of powers
zeta	A vector of zeta values

plot.fitted_distribution
Plot a fitted distributions object

Description

Creates a plot for a result of ‘PCNMA::fit_distribution’.

Usage

```
## S3 method for class 'fitted_distribution'
plot(
  fit,
  CI = FALSE,
  km = FALSE,
  km_alpha = 1,
  linewidth = 0.75,
  linetype = "dashed",
  theme = "bw",
  facet_by = "Treatment",
  ...
)
```

Arguments

fit	A 'PCNMA::fitted_distribution' object
CI	Include a confidence interval?
km	Add the original KM curve?
...	For S3 consistency

plot.fitted_model *Plots for an NMA modelr*

Description

Plots for an NMA modelr

Usage

```
## S3 method for class 'fitted_model'
plot(
  model,
  type = "trace",
  pars = parsForStan,
  prob = 0.95,
  ordered = FALSE,
  xLims = NULL
)
```

Arguments

model	A [PCNMA::fitted_model] object
type	Type of plot to produce

plot.km_obj *Plot a KM curve*

Description

Plot a KM curve

Usage

```
## S3 method for class 'km_obj'  
plot(  
  fit,  
  type = "survival",  
  risk.table = TRUE,  
  break.x.by = 5,  
  xMax = 40,  
  risktable.height = 0.3,  
  ...  
)
```

Arguments

fit	A 'PCNMA::km_obj' object
type	See 'ggsurvfit::ggsurvfit' for details.
risk.table	Add numbers at risk?
break.x.by	A numeric value for splitting x axis
xMax	The maximum time value to plot
risktable.height	The proportion of the figure to be taken up by the risk table
...	For S3 consistency

Value

A plotted km curve

plot_network	<i>Plot a network of evidence</i>
--------------	-----------------------------------

Description

Plot a network of evidence

Usage

```
plot_network(network, ...)
```

Arguments

network	A network dataset
---------	-------------------

rmst	<i>Hazard ratios of a fitted model</i>
-------------	--

Description

Hazard ratios of a fitted model

Usage

```
rmst(fit, x, ...)
```

Arguments

fit	A [PCNMA::fitted_distribution] object
x	Time to calculate RMST for
...	For S3 consistency

summary.fitted_distribution	<i>Summary of a set of fitted models</i>
------------------------------------	--

Description

Summary of a set of fitted models

Usage

```
## S3 method for class 'fitted_distribution'
summary(fit, AIC = FALSE, median = FALSE)
```

Arguments

fit	A [PCNMA::fitted_distribution] object.
AIC	Returns the AIC scores for a set of models
median	Returns a table of median estimates for a set of models

summary.fitted_model *Summary of an NMA model*

Description

Summary of an NMA model

Usage

```
## S3 method for class 'fitted_model'  
summary(model, likelihood, effect)
```

Arguments

model A [PCNMA::fitted_model] object

summary.hr_obj *Summarise a Hazard Ratio*

Description

Summarise a Hazard Ratio

Usage

```
## S3 method for class 'hr_obj'  
summary(hr)
```

summary.km_obj *Summarise KM data*

Description

Summarise KM data

Usage

```
## S3 method for class 'km_obj'  
summary(fit, ...)
```

Arguments

fit A 'PCNMA::km_obj' object

Value

A summary table of the KM data

Index

.fit_distribution, 2
boxTid, 2
coef.fitted_distribution, 3
fit_distribution, 3
fit_model, 4
gen_network, 4
gen_network_data, 5
H, 5
hr, 6
km_estimates, 6
phi, 7
plot.fitted_distribution, 7
plot.fitted_model, 8
plot.km_obj, 8
plot_network, 9
rmst, 10
summary.fitted_distribution, 10
summary.fitted_model, 11
summary.hr_obj, 11
summary.km_obj, 11

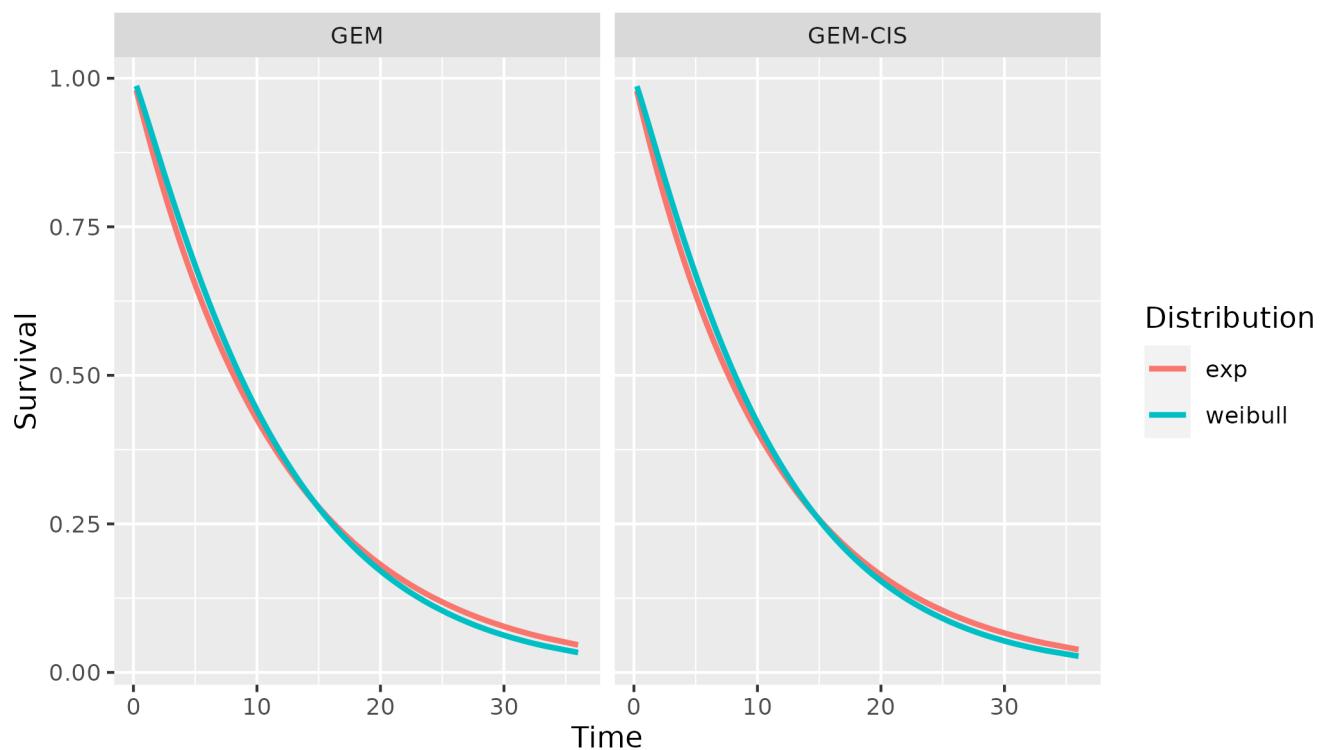
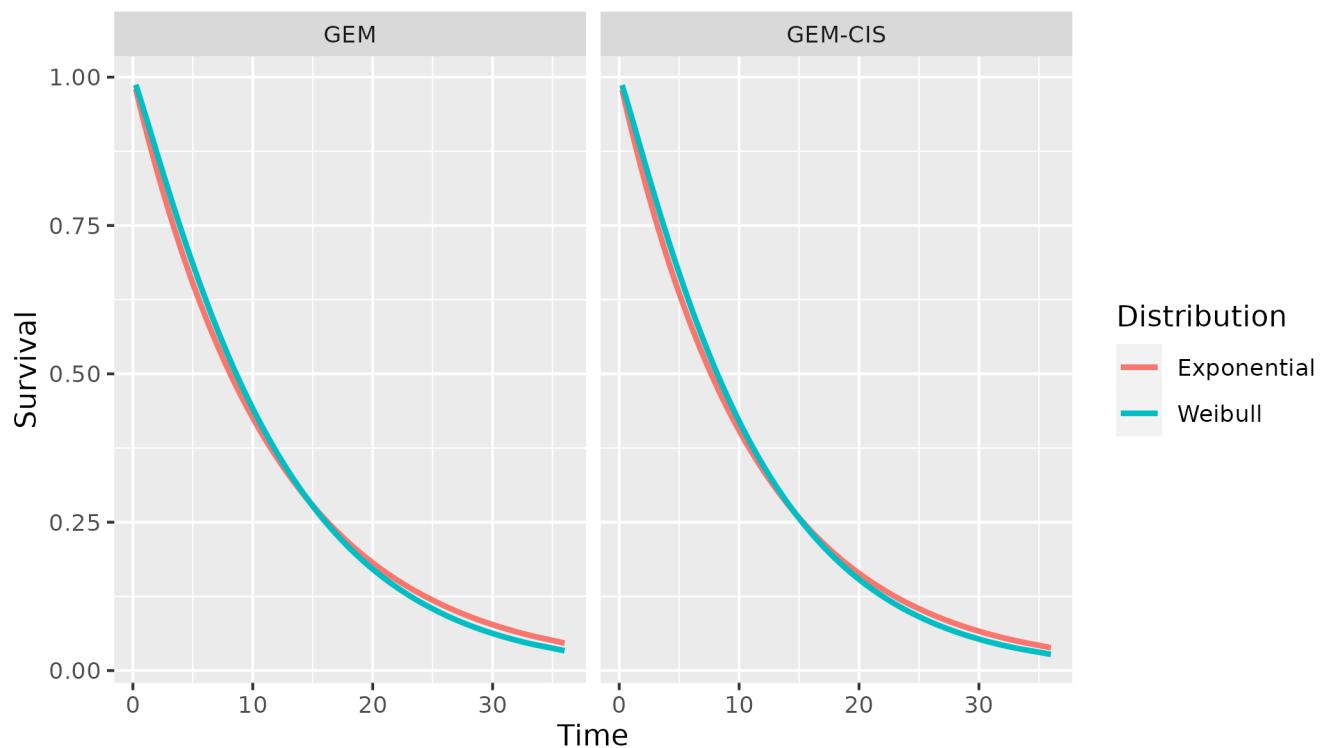


Figure C.1: Fitted Models as in Flexsurv

Figure C.2: Fitted Models with *fit_distribution*