

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

Matthew Knowles



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

Supervisor: Dr. Kate Ren

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

Abstract

Gemcitabine (GEM) in combination with nab-Paclitaxel (NAB) or Capecitabine (CAP) provides improved median OS and restricted mean survival time (RMST) estimates for the treatment of advanced/metastatic pancreatic cancer.

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.

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

CONTENTS

1	Introduction	6
1.1	Pancreatic Cancer	6
1.2	Treatment Landscape	6
1.3	Project Aim	7
2	Survival Analysis Bakground	8
2.1	Background and Survival Functions	8
2.2	Regression Models for Survival	8
2.2.1	Accelerated Failure Time (AFT) Models	8
2.2.2	Proportional Hazards (PH) Models	9
2.3	Key Survival Metrics	9
2.3.1	The Hazard Ratio	9
2.3.2	Restriced Mean Survival Time	10
2.4	Parametric Models for Survival Analysis	10
2.4.1	Model Setup	10
2.4.2	Fitting Models	11
2.4.2.1	Right Censoring	11
2.4.2.2	Interval Censoring	11
3	Network Meta Analysis Theory	12
3.1	Building a Network of Evidence	12
3.2	Standard NMA Model	13
3.3	Multi-Level Network Meta Regression	13
3.4	Survival ML-NMR	14
4	Studies	15
5	Survival Analysis of Pancreatic Cancer Trials	19
6	NMA of Pancreatic Cancer Trials	20
6.1	Network of Evidence	20
6.2	Model Fitting and Selection	21
6.3	Results	21
7	Conclusion and Discussion	26
7.1	Conclusion	26
7.2	Discussion and Further Work	26
A	Survival Analysis Results	29
A.1	Parametric Models	29

B Additional NMA Results	34
B.1 Model Convergence	34
C The PCNMA Package	36
C.1 Survival Functions	36

LIST OF FIGURES

3.1	Visualisation of combining trials into a network of evidence	13
4.1	Forest plot for median OS of GEM in each study	16
4.2	Forest plot for median OS of the comparator in each study	17
4.3	KM curves for each study	18
6.1	Network of evidence	20
6.2	Survival of each treatment in each population	22
6.3	Hazards of each treatment in each population	23
6.4	RMST of each treatment in each population	24
6.5	Median of each treatment in each population	25
A.1	Cunningham (2009) Parametric extrapolations	30
A.2	Goldstein (2015) Parametric extrapolations	30
A.3	Goncalves (2012) Parametric extrapolations	31
A.4	Kindler (2011) Parametric extrapolations	31
A.5	Oettle (2005) Parametric extrapolations	32
A.6	Rocha Lima (2004) Parametric extrapolations	32
A.7	Spano (2008) Parametric extrapolations	33
B.1	Trace plot for the FE log-normal ML-NMR model	35
C.1	Fitted Models as in Flexsurv	49
C.2	Fitted Models with <i>fit_distribution</i>	49

INTRODUCTION

1.1 Pancreatic Cancer

Pancreatic cancer is the 10th most common cancer in the UK, accounting for 3% of all new cases [Cancer Research UK, 2024]. Pancreatic cancer has a particularly poor prognosis. In 2016-2018, there were 9,558 deaths from 10,452 cases. Pancreatic cancer is hard to detect at early stages, meaning most people who present with symptoms already have advanced-stage pancreatic cancer. 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].

Risk factors of pancreatic cancer include smoking, diabetes, obesity, and high-fat diets. Smoking is the dominant cause, 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].

1.2 Treatment Landscape

Gemcitabine (GEM) ($C_9H_{11}F_2N_3O_4$) is a standard first-line treatment for pancreatic cancer administered intravenously [for Health and Excellence, 2018]. 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 (SOR) ($C_{21}H_{16}ClF_3N_4O_3$), nab-Paclitaxel (NAB)¹ and Irinotecan (IRI) ($C_{33}H_{38}N_4O_6$).

A Network Meta Analysis (NMA) conducted by [Gresham et al., 2014], which included GEM, GEM-NAB, and GEM-CAP, found all three to be associated with statistically significant improvements in Overall Survival (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, defined as in Equation 1.1.

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

$$\frac{\frac{\text{GEM Median OS}}{HR} - \text{GEM Median OS}}{\frac{\text{GEM Median PFS}}{HR} - \text{GEM Median PFS}} \quad (1.1)$$

1.3 Project Aim

To assess the efficacy of GEM alone versus GEM in combination with one of the six aforementioned treatments for the treatment of advanced/metastatic pancreatic cancer, a modified Multi-Level Network Meta Regression (ML-NMR) was used. The focus of this project was not on a thorough literature review, but on the methodology, so inclusion criteria were not particularly strict, but studies had to be Phase III, 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.

SURVIVAL ANALYSIS BAKGROUND

2.1 Background and 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 .

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

Definition 2.1.2: Hazard Function

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

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

2.2 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.

2.2.1 Accelerated Failure Time (AFT) Models

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

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

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

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

This is equivalent to defining a random variable T such that

$$T = T_0/\Psi(\mathbf{x}).$$

Here, T_0 has survivor function S_0 . It is required that $\Psi(\mathbf{x}) \geq \mathbf{0}$ and $\Psi(\mathbf{0}) = \mathbf{1}$ [[NOTE TO MATT: WHY?]], leading to the natural choice

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

We can then write

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

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

2.2.2 Proportional Hazards (PH) 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

$$h(t, \mathbf{x}) = \exp(\beta' \mathbf{x}) \mathbf{h}_0(\mathbf{t}) \quad (2.1)$$

Consider the following definition.

Definition 2.2.1: 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.

2.3 Key Survival Metrics

2.3.1 The Hazard Ratio

The Hazard Ratio (HR) follows from Equation 2.1. Consider two treatments, $i = 1, 2$, then $h_1(t, \mathbf{x}) = \exp(\beta' \mathbf{x}) \mathbf{h}_0(\mathbf{t})$ and $h_2(t, \mathbf{x}) = \exp(\beta' \mathbf{x}) \mathbf{h}_0(\mathbf{t})$. The HR is obtained as in Equation 2.2.

$$HR = \frac{h_1}{h_2} = \exp(\beta' \mathbf{x}) \quad (2.2)$$

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 2.3.1: 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.

2.3.2 Restricted Mean Survival Time

The Restricted Mean Survival Time (RMST) is 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. We therefore define it formally as

Definition 2.3.2: RMST

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

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

2.4 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 National Institute for Health and Care Excellence (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.

2.4.1 Model Setup

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

$$f(t|\mu(\mathbf{z}), \alpha(\mathbf{z})). \quad (2.3)$$

Equation 2.3 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

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

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

$$\begin{aligned} g_0(\mu(\mathbf{z})) &= \gamma_0 + \beta_0^T \mathbf{z} \\ g_r(\alpha_r(\mathbf{z})) &= \gamma_r + \beta_\gamma^T \mathbf{z} \end{aligned} \quad (2.4)$$

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

Flexsurv contains several built-in models. Appendix ?? presents these models.

2.4.2 Fitting Models

Let $t_i, i \in \{1, \dots, n\}$, be a sample of times from n individuals. Define c_i such that

$$c_i = \begin{cases} 1 & \text{if } t_i \text{ is an observed death time} \\ 0 & \text{if } t_i \text{ is censored} \end{cases}.$$

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.

2.4.2.1 Right Censoring

In the case of right-censoring and nothing else, the likelihood for the parameters $\theta = \{\gamma, \beta\}$ required in Equation 2.4 is given by

$$l(\theta | \mathbf{t}, \mathbf{c}, \mathbf{s}) = \frac{\prod_{i:c_i=1} f_i(t_i) \prod_{i:c_i=0} S_i(t_i)}{\prod_i S_i(s_i)} \quad (2.5)$$

2.4.2.2 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

$$l(\theta | \mathbf{t}^{\min}, \mathbf{t}^{\max}, \mathbf{c}, \mathbf{s}) = \frac{\prod_{i:c_i=1} f_i(t_i) \prod_{i:c_i=0} (S_i(t_i^{\min}) - S_i(t_i^{\max}))}{\prod_i S_i(s_i)} \quad (2.6)$$

Maximum Likelihood Estimation is performed in **R** using the analytic derivatives of Equation 2.5 and/or Equation 2.6.

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 becomes

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

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 a network meta analysis. This process expands to trials that compare more than two treatments by weighting the edges by the number of trials making that particular comparison.

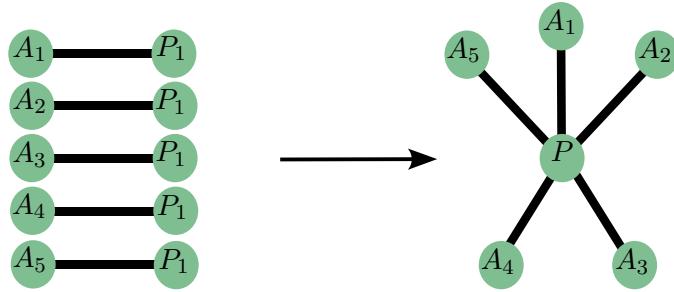


Figure 3.1: Visualisation of combining trials into a network of evidence

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 , the standard NMA model is written as in Equation ???. This summary outcome may be, for example, HRs, or RMST values. In Equation 3.1, 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 ???, μ_j and δ_{jk} are study-specific intercepts and study-specific relative effect of treatment k versus the reference treatment. Together, Equation 3.1 and Equation 3.2 give the standard NMA model.

$$y_{jk} \sim \pi_{Agg}(\theta_{jk}) \quad (3.1)$$

$$g(\theta_{jk}) = \mu_j + \delta_{jk} \quad (3.2)$$

There are two types of NMA: fixed effect (FE) and relative effects (RE). 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 Multi-Level Network-Meta-Regression (ML-NMR) have sought to relax this assumption by using IPD from at least one of the studies in a population.

3.3 Multi-Level Network Meta Regression

The derivation in this chapter is based on the work of ??, but modified under the context of this project. Under an NMA framework, we have J RCTs investigating a subset K_j ($j = 1, \dots, J$) of K treatments. In this project, $|K_j| = 2 \forall j$. Depending on data availability, we may have IPD for some studies, and only aggregate data for the remaining.

Definition 3.3.1: General IPD Meta-Regression Model

Let y_{ijk} be the IPD outcome for individual $i = 1, \dots, N_{kj}$ 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 $\eta_{jk}(\mathbf{x}_{ijk})$. 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 3.1-3.2. 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 3.3, the individual marginal likelihood function.

$$L_{ijk}^{\text{Mar}}(\xi; y_{ijk}) = \int_{\mathfrak{X}} L_{ijk|x}^{\text{Con}}(\xi; y_{ijk}, x) dx \quad (3.3)$$

It is clear from Equation 3.3 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 don't 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 3.3.

It is likely that a closed-form of Equation 3.3 does not exist. We can therefore take a set of N integration points, \hat{x} from f_{jk} , giving Equation 3.4.

$$L_{ijk}^{\text{Mar}}(\xi; y_{ijk}) \approx \frac{1}{N} \sum_{\hat{x}} L_{ijk|x}^{\text{Con}}(\xi; y_{ijk}, \hat{x}) \quad (3.4)$$

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} . Let $S_{jk}(t|x)$ 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

$$L_{ijk|x}^{\text{Con}}(\zeta; t_{ijk}, c_{ijk}, x_{ijk}) = S_{jk}(t_{ijk}|x_{ijk}) h_{jk}(t_{ijk}|x_{ijk})^{c_{ijk}} \quad (3.5)$$

The forms of S and h depend on the specific survival models chosen.

CHAPTER
FOUR

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.

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 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.

Study	Treatment	N	Median Age	Proportion Male	Median OS (Months)
Cunningham 2009	GEM	266	62.0	0.580	6.2 (5.5, 7.2)
Cunningham 2009	GEM-CAP	267	62.0	0.570	7.1 (6.2, 7.8)
Goldstein 2015	GEM	430	63.0	0.600	6.6 (6.0, 7.2)
Goldstein 2015	GEM-NAB	431	62.0	0.570	8.7 (7.9, 9.7)
Goncalves 2012	GEM	52	64.0	0.620	9.2 (7.7, 11.6)
Goncalves 2012	GEM-SOR	52	61.0	0.580	8.0 (6.0, 10.8)
Kindler 2011	GEM	316	61.0	0.590	8.3 (6.9, 10.3)
Kindler 2011	GEM-AXI	314	61.0	0.610	8.5 (6.9, 9.5)
Oettle 2005	GEM	282	63.0	0.535	6.3 (5.4, 6.9)
Oettle 2005	GEM-PEM	283	63.0	0.604	6.2 (5.4, 6.9)
Rocha Lima 2004	GEM	180	60.2	0.533	6.6 (5.2, 7.8)
Rocha Lima 2004	GEM-IRI	180	63.2	0.572	6.3 (4.7, 7.5)
Spano 2008	GEM	34	61.0	0.470	5.6 (3.9, 8.8)
Spano 2008	GEM-AXI	69	65.0	0.510	6.9 (5.3, 10.1)

Table 4.1: Included studies

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.

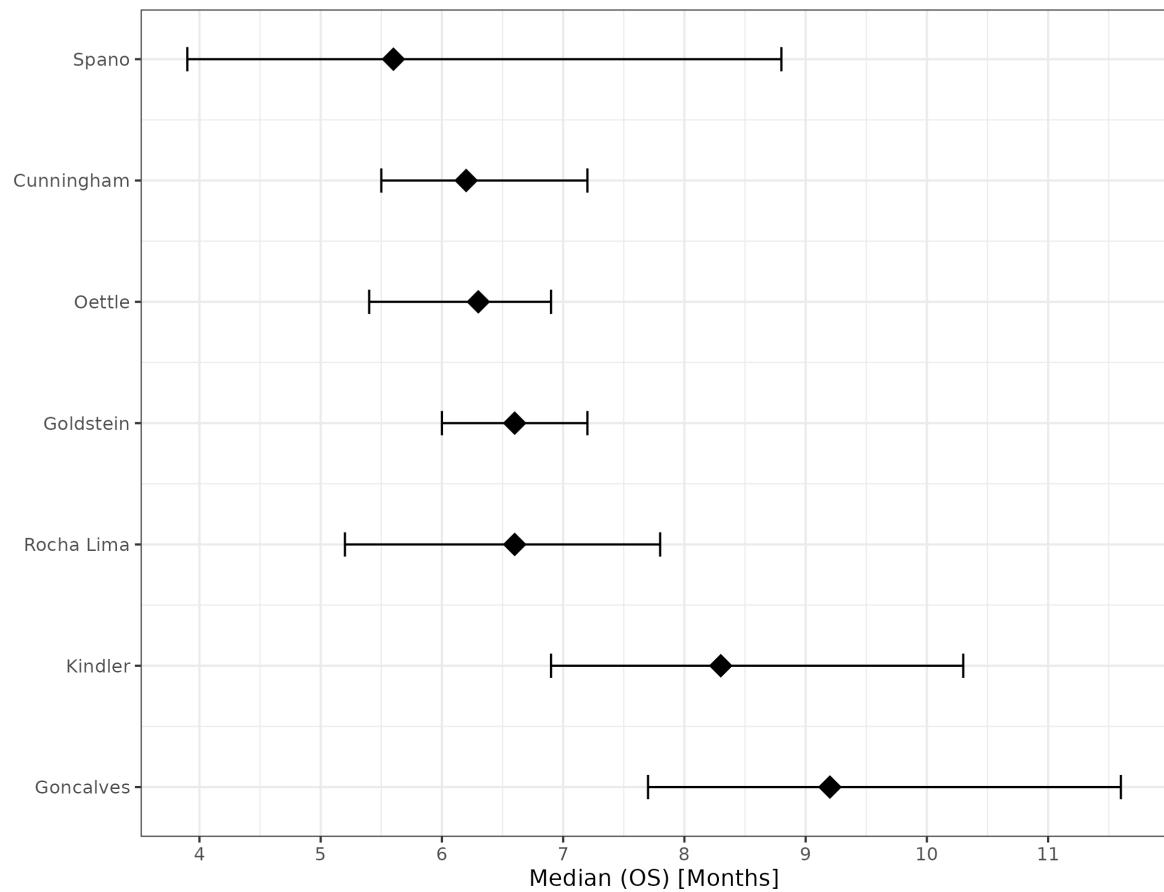


Figure 4.1: Forest plot for median OS of GEM in each study

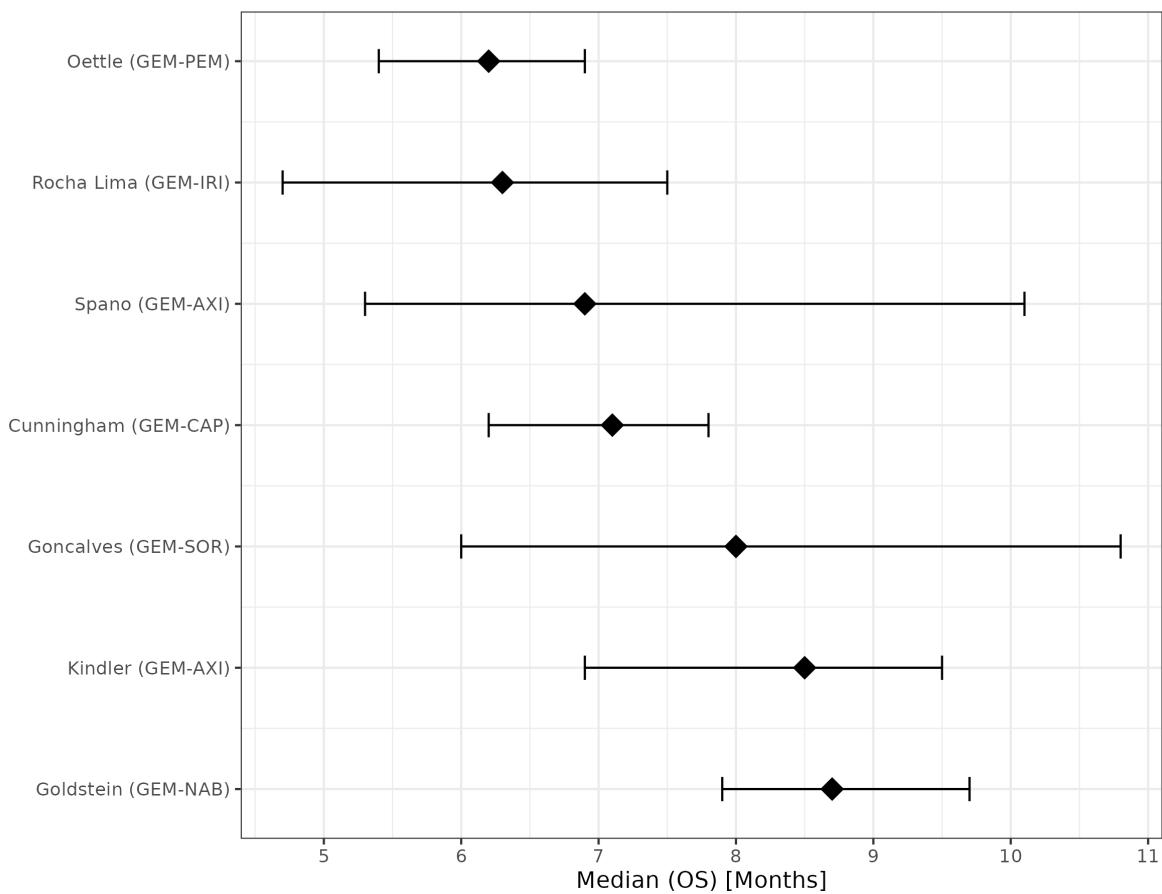


Figure 4.2: Forest plot for median OS of the comparator in each study

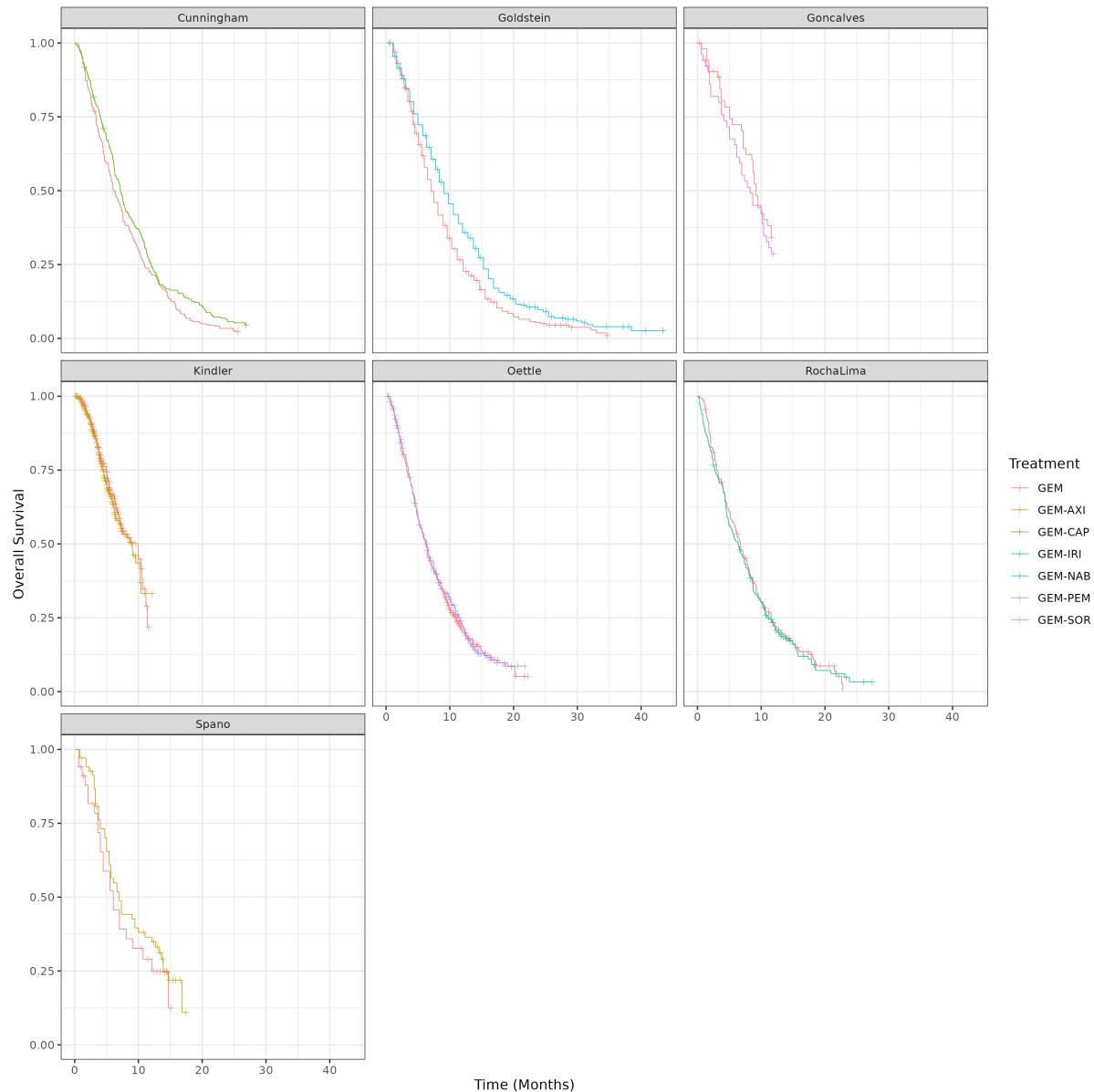


Figure 4.3: KM curves for each study

CHAPTER
FIVE

SURVIVAL ANALYSIS OF PANCREATIC CANCER TRIALS

NMA OF PANCREATIC CANCER TRIALS

6.1 Network of Evidence

Figure 6.1 presents the network of evidence for this NMA. Only aggregate data (AgD) was available for each study. 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.

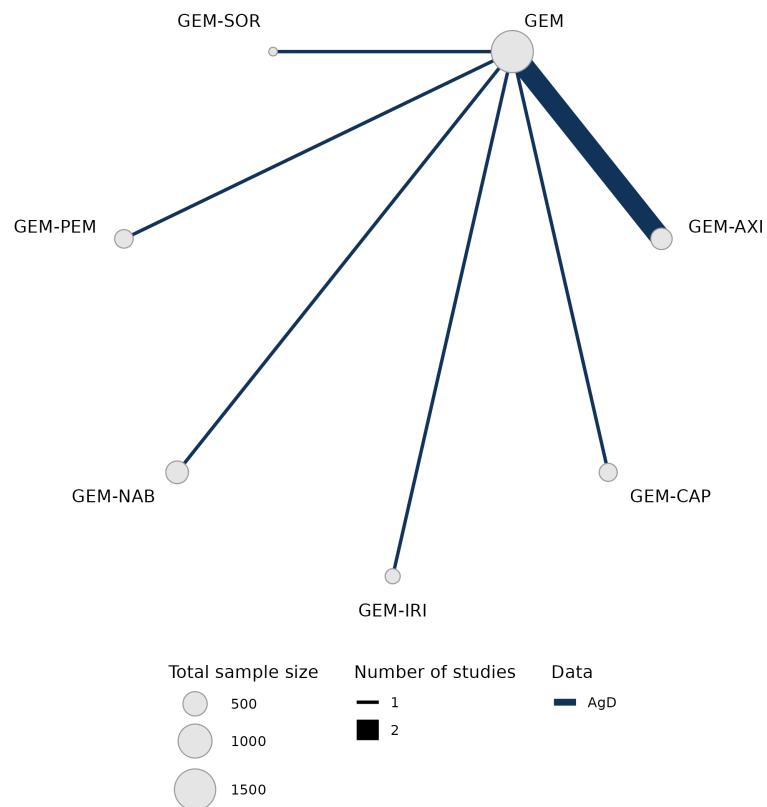


Figure 6.1: Network of evidence

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*.

Definition 6.2.1: Half Normal Distribution

Let X be a normal distribution such that $X \sim N(0, \sigma^2)$. Then $Y = |X|$ follows a half-normal distribution. In particular, the half-normal distribution has PDF

$$f(x, \sigma) = \frac{\sqrt{2}}{\sigma\sqrt{\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right)$$

With $x \in [0, \infty)$, and $\sigma > 0$.

For each model, sampling was done using 2000 iterations on four chains. The first 1000 iterations were warmup iterations.

Table 6.1 presents the selection statistics for each model. 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 B, Figure B.1, and indicated good convergence due to consistent peaks and troughs across each treatment arm. The fixed effect model was also deemed to be clinically appropriate due to homogeneity in the patient population.

Likelihood	Effect	DIC	LOOIC
Gamma	Fixed	NA	NA
Gamma	Random	Inf	Inf
Gen Gamma	Fixed	NA	NA
Gen Gamma	Random	NA	NA
Gompertz	Fixed	15388.1764	15381.1184
Gompertz	Random	Inf	Inf
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 6.2 and Figure 6.3 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 followest 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.

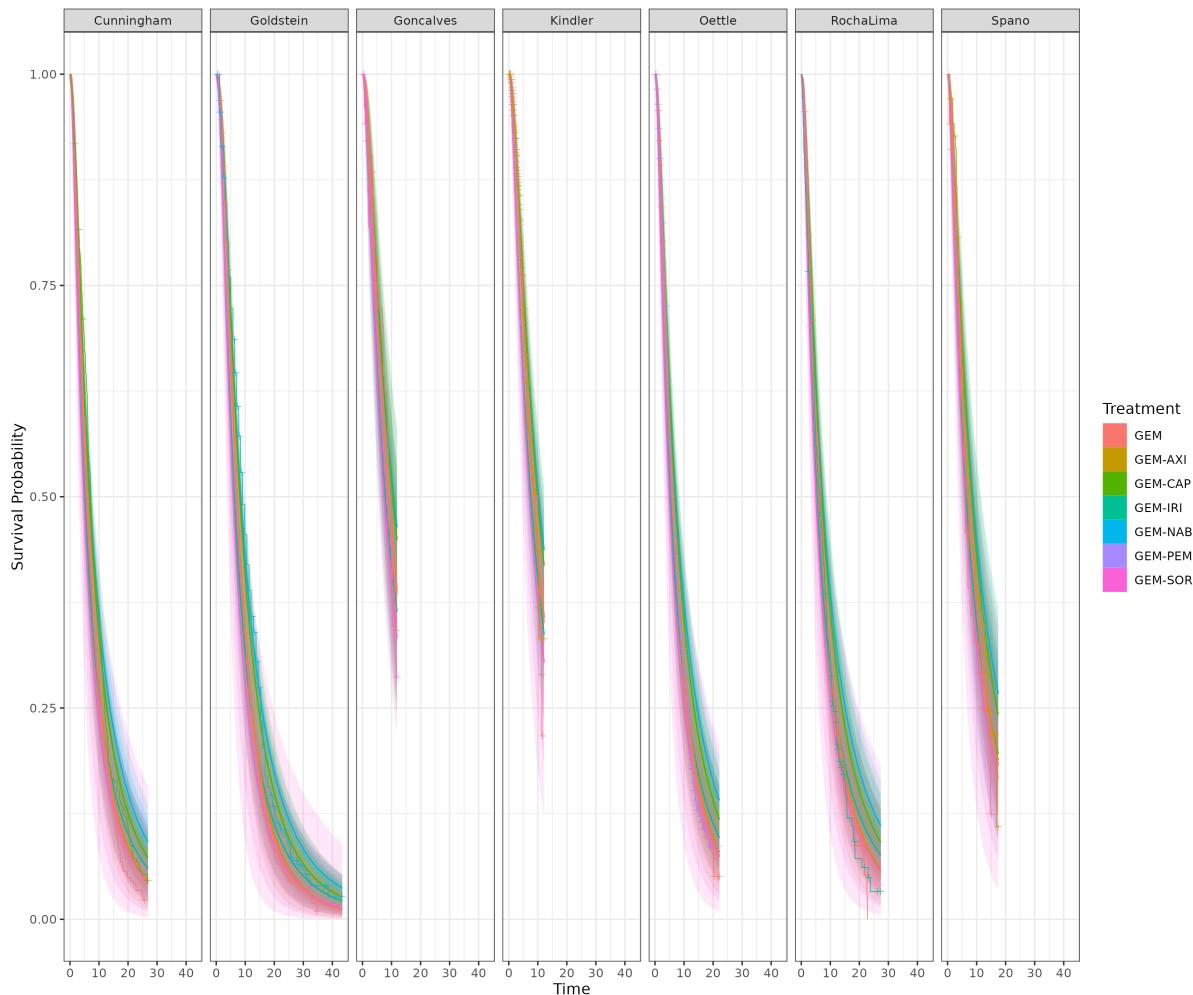


Figure 6.2: Survival of each treatment in each population

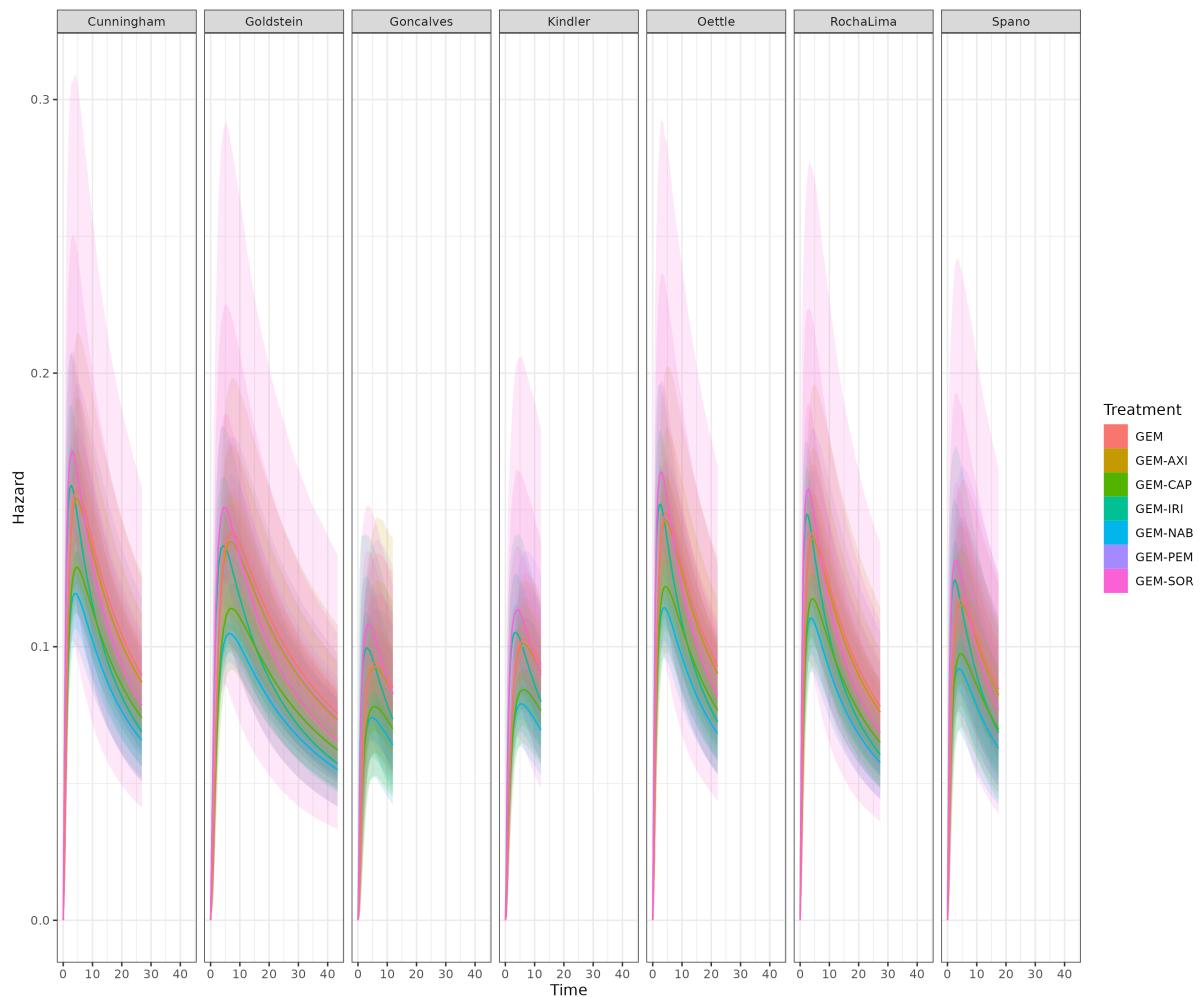


Figure 6.3: Hazards of each treatment in each population

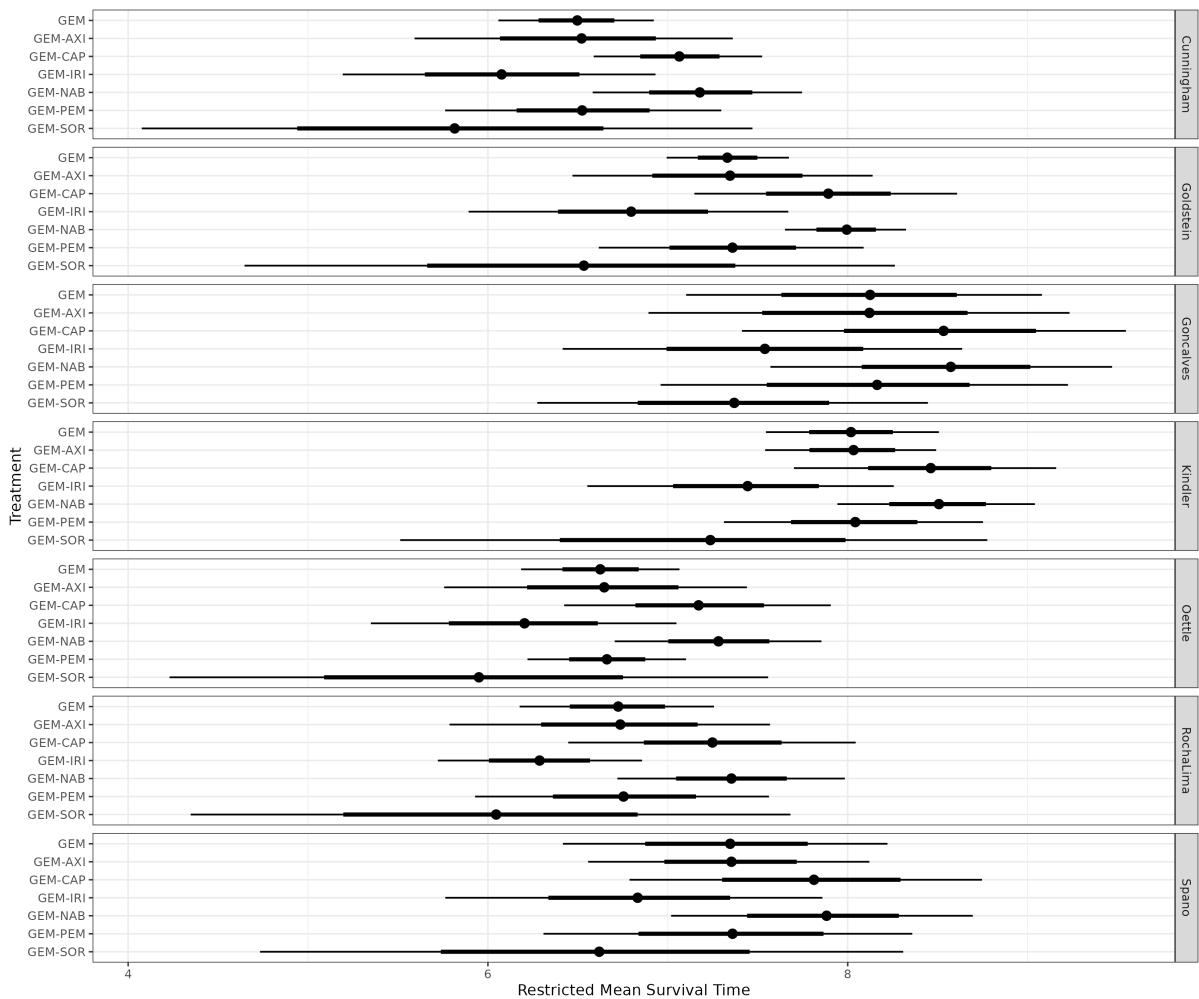


Figure 6.4: RMST of each treatment in each population

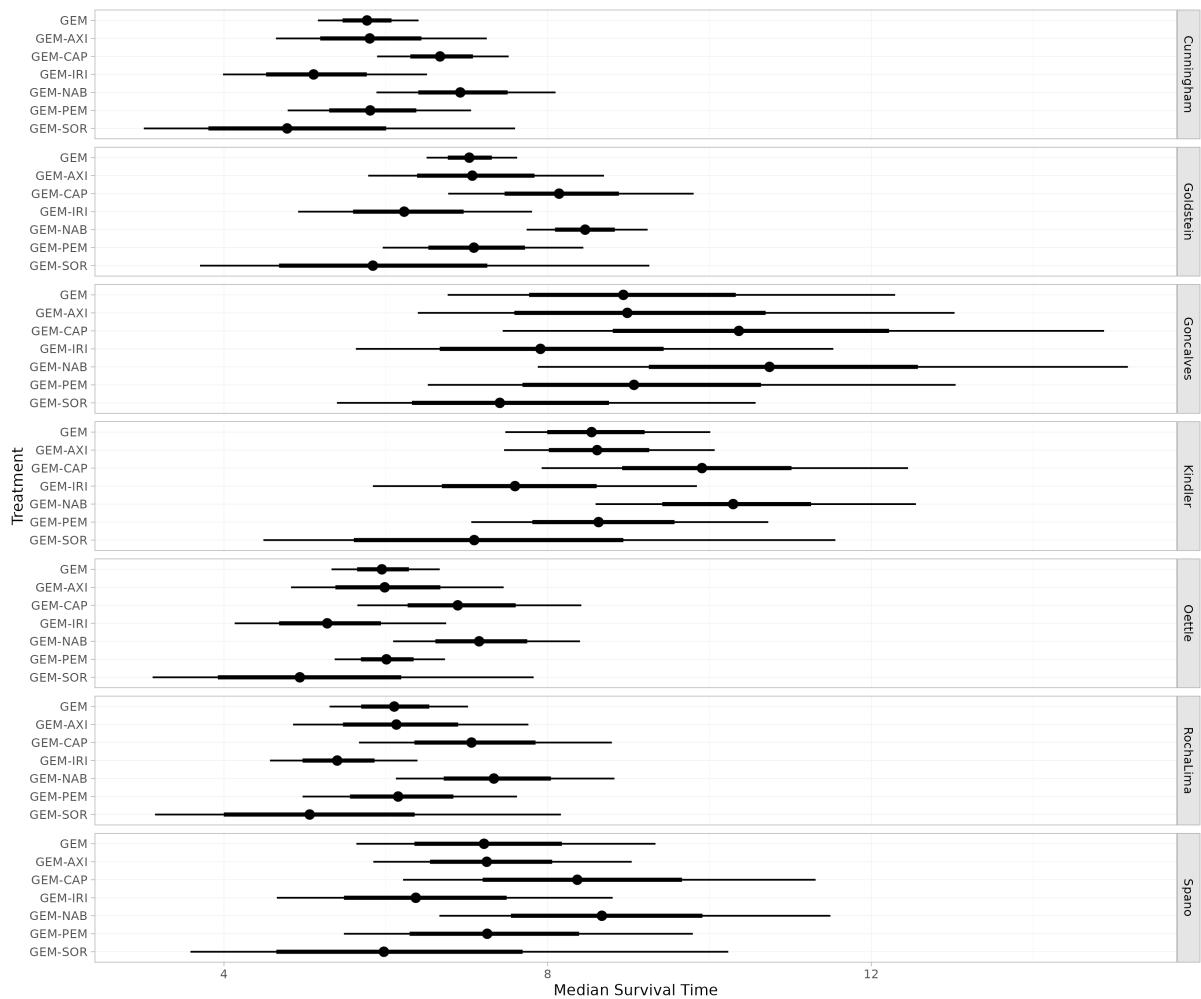


Figure 6.5: Median of each treatment in each population

CONCLUSION AND DISCUSSION

7.1 Conclusion

GEM-NAB and GEM-CAP offer a better alternative to GEM monotherapy for the treatment of advanced/metastatic pancreatic cancer in terms of median OS and RMST. GEM-AXI and GEM-PEM do not significantly improve median OS or RMST compared to GEM monotherapy, and GEM-IRI and GEM-SOR both give worse median OS and RMST compared to GEM alone.

7.2 Discussion and Further Work

The ML-NMR framework is ideal for working with a combination of both IPD and AgD. Unfortunately, no IPD could be obtained for this dissertation, but would lead to interesting future 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 solidfy the place of GEM-NAB and GEM-CAP as the best treatment options for advanced/pancreatic cancer.

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.

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.
- [Cox, 1972] Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2):187–202.
- [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.
- [for Health and Excellence, 2018] for Health, N. I. and Excellence, C. (2018). Pancreatic cancer in adults: diagnosis and management.
- [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.
- [Jackson, 2016] Jackson, C. H. (2016). flexsurv: a platform for parametric survival modeling in r. *Journal of statistical software*, 70.
- [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.
- [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*.
- [Pancreatic Cancer UK, 2022] Pancreatic Cancer UK (2022). Signs and symptoms of pancreatic cancer.
- [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.

Appendices

APPENDIX

A

SURVIVAL ANALYSIS RESULTS

A.1 Parametric Models

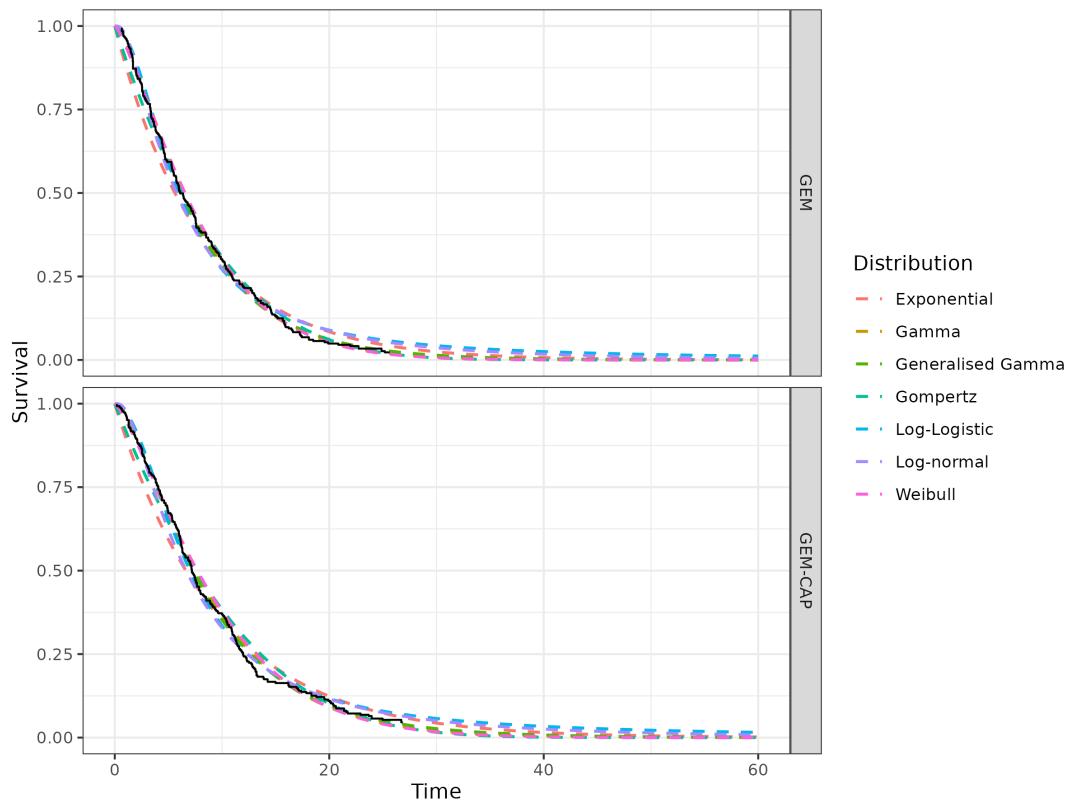


Figure A.1: Cunningham (2009) Parametric extrapolations

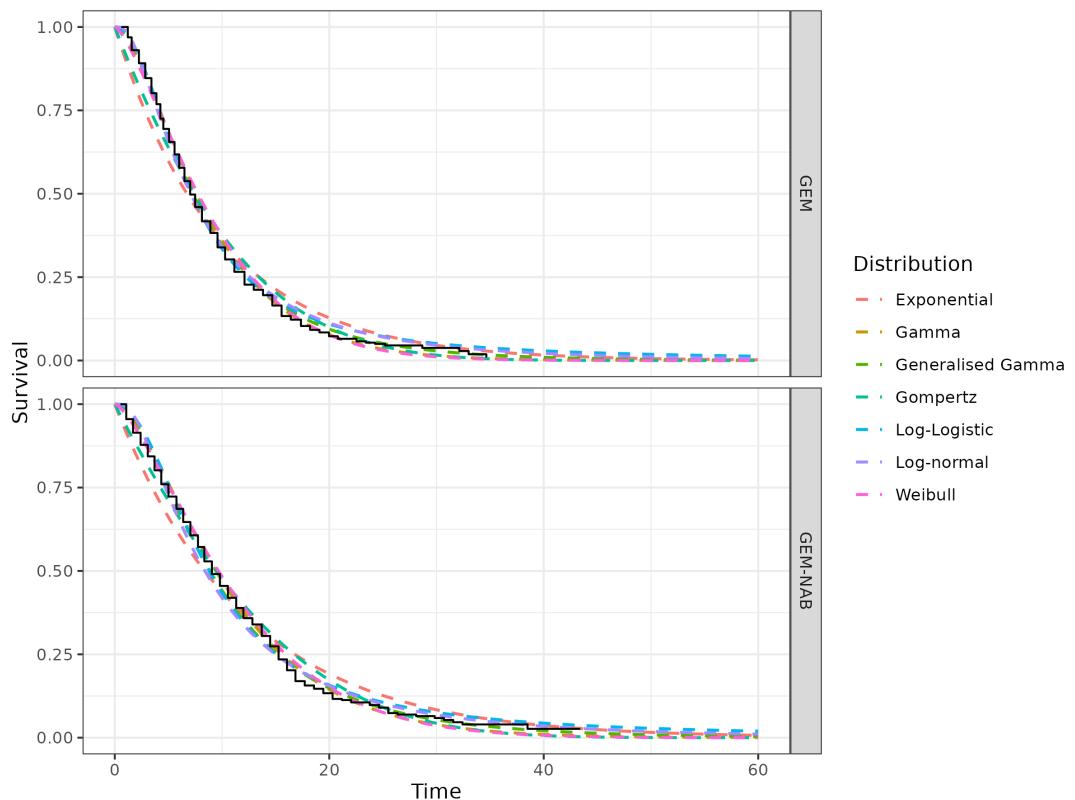


Figure A.2: Goldstein (2015) Parametric extrapolations

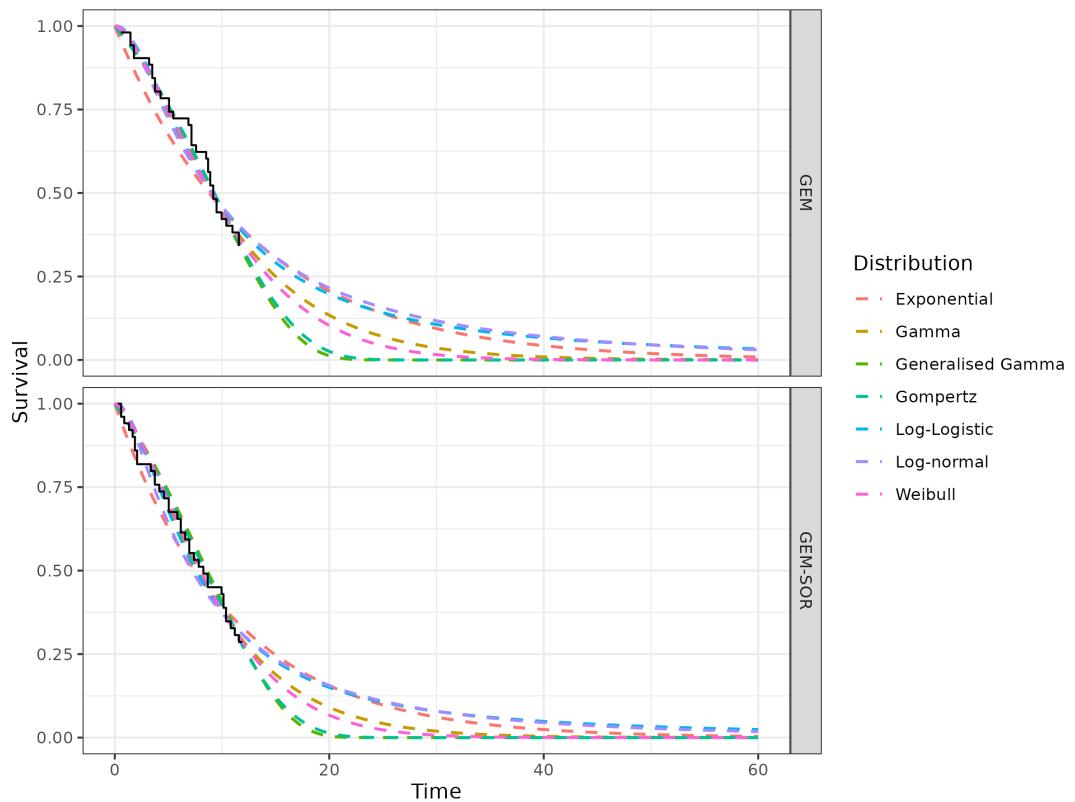


Figure A.3: Goncalves (2012) Parametric extrapolations

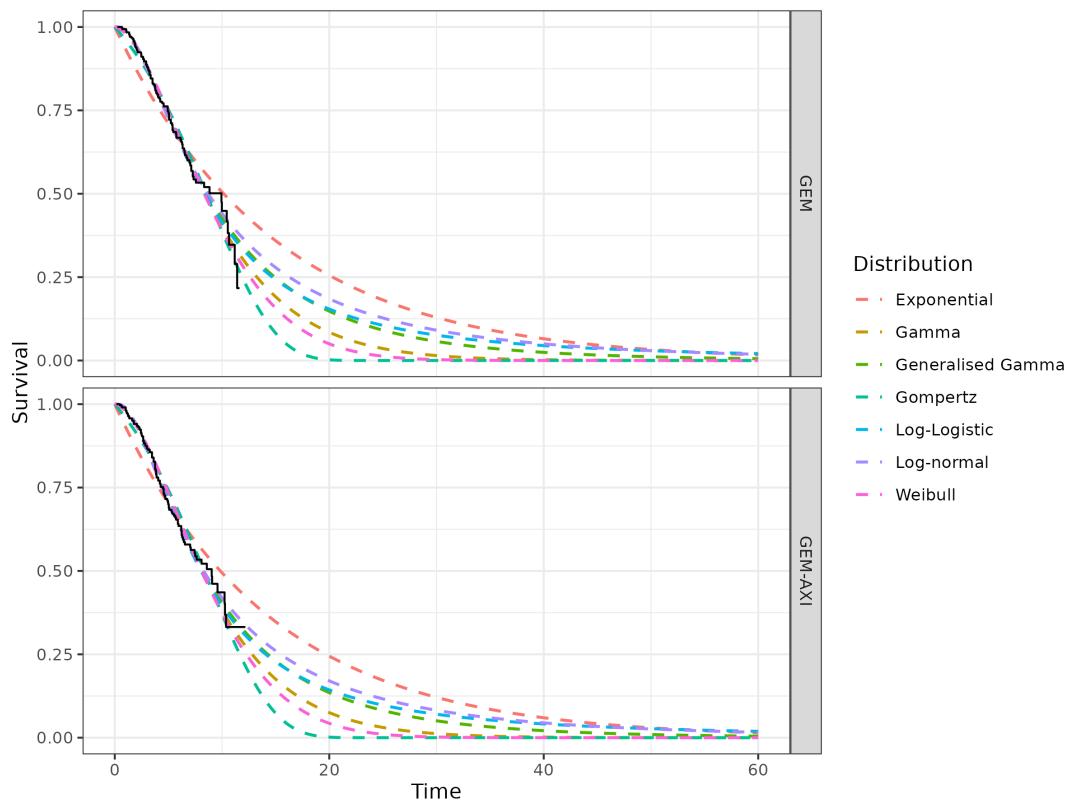


Figure A.4: Kindler (2011) Parametric extrapolations

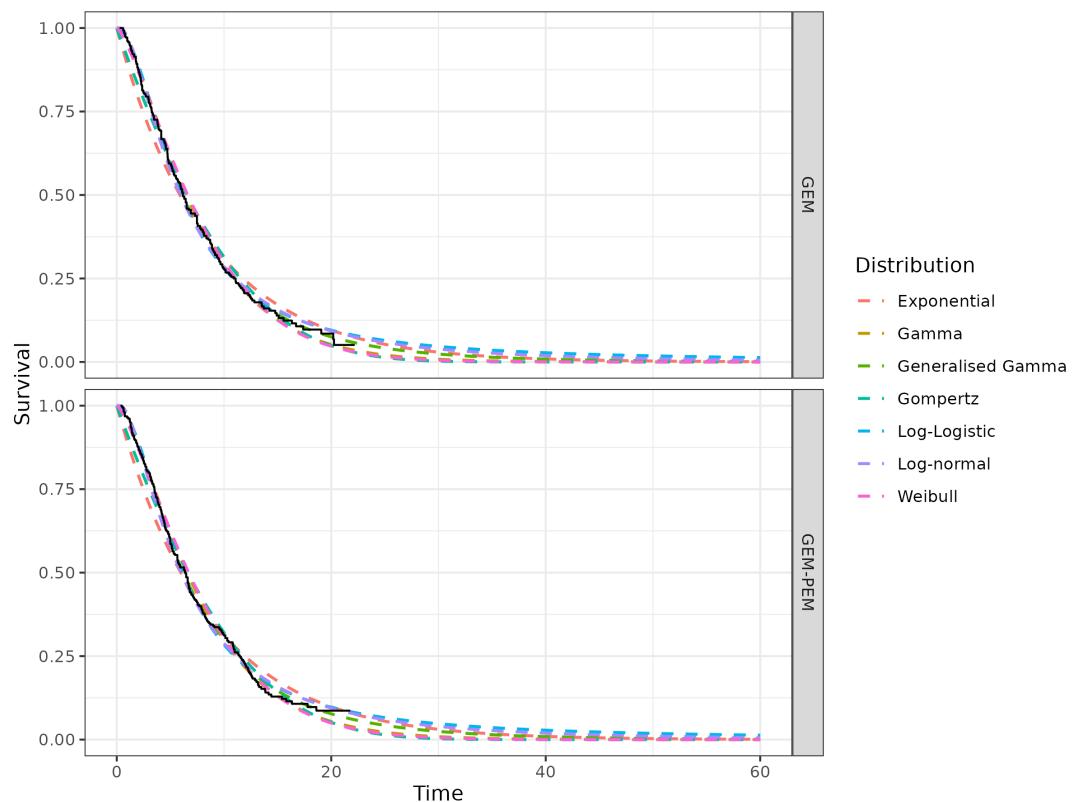


Figure A.5: Oettle (2005) Parametric extrapolations

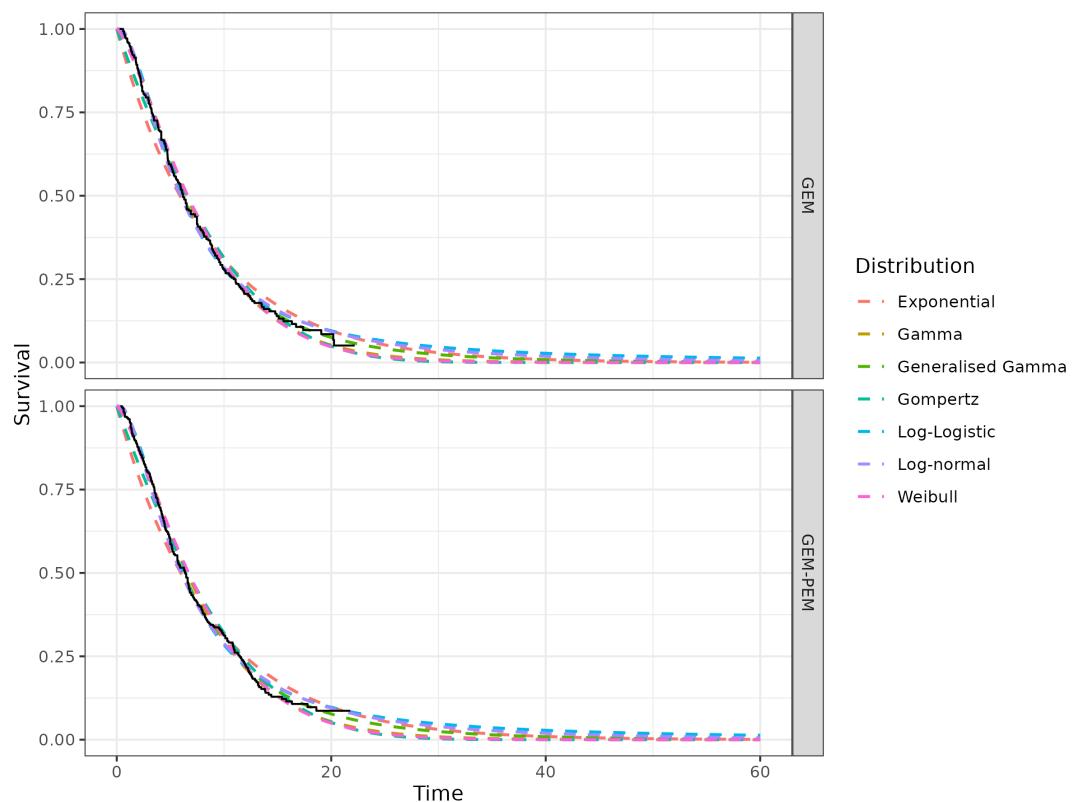


Figure A.6: Rocha Lima (2004) Parametric extrapolations

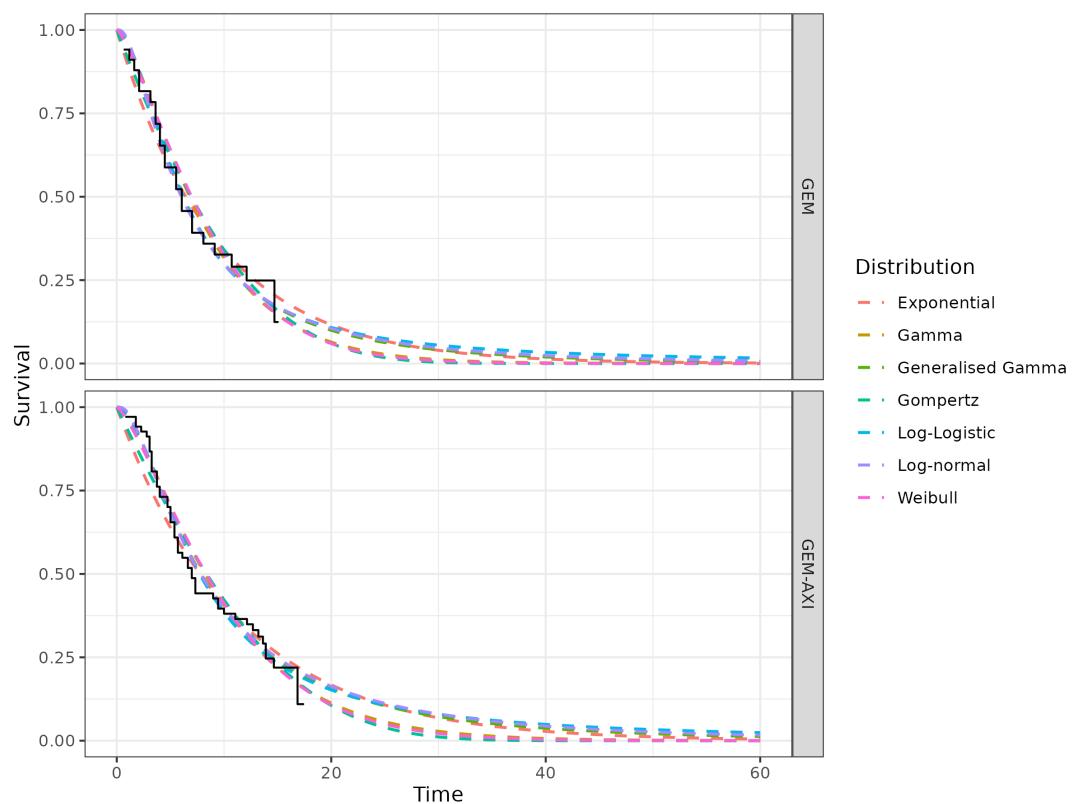


Figure A.7: Spano (2008) Parametric extrapolations

APPENDIX
B

ADDITIONAL NMA RESULTS

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

B.1 Model Convergence

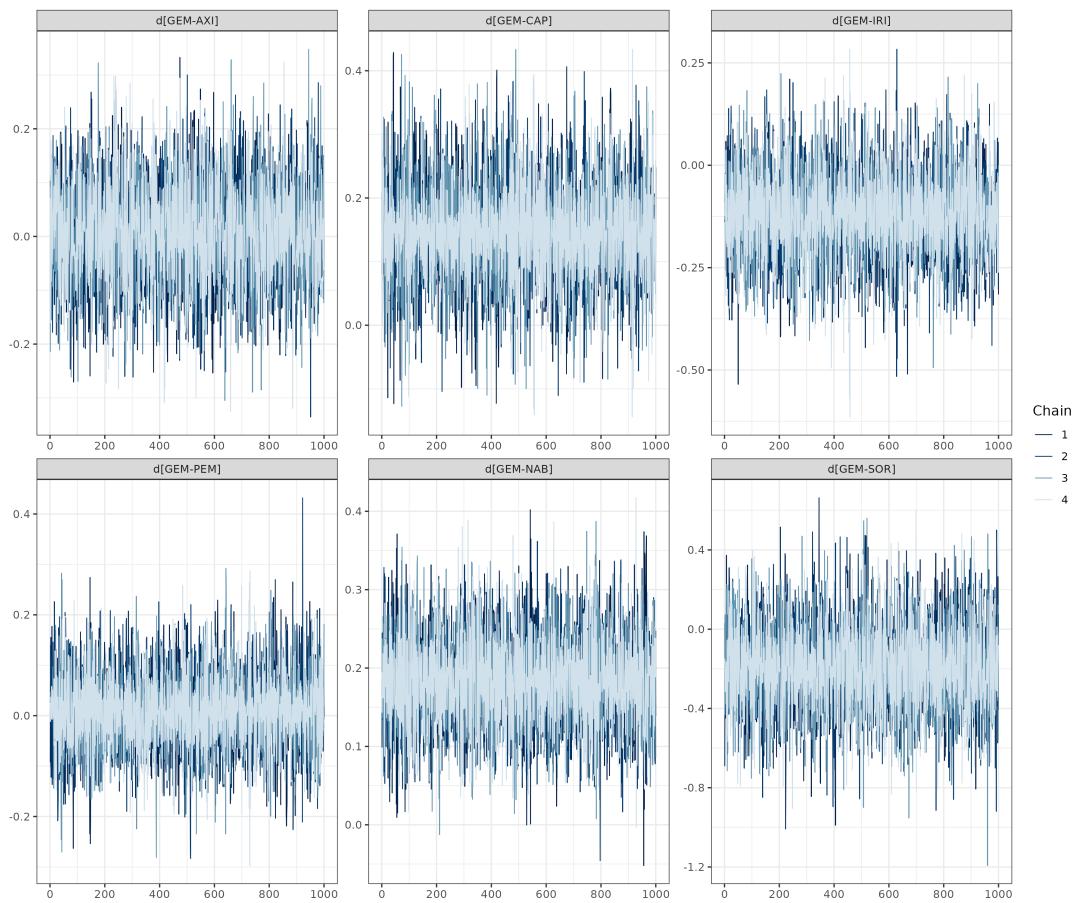


Figure B.1: Trace plot for the FE log-normal ML-NMR model

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 are somewhat untidy. 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 **purrr::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.

Package ‘PancSurv’

March 10, 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

RoxxygenNote 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_aic	9
plot_network	9
save_nma_summary	10
summary.fitted_distribution	10
summary.hr_obj	11
summary.km_obj	11

Index**12**

.fit_distribution *Fit a single survival distribution*

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

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

Value

A [flexsurv::flexsurvreg] object

boxTid *Box-Tidwell Transformation*

Description

Box-Tidwell Transformation

Usage

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

Arguments

x	A real value
p	The p-value to raise x to

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

Description

Coefficients of fitted models

Usage

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

Arguments

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

```
fit_distribution      Fit survival distributions to a dataset.
```

Description

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

Usage

```
fit_distribution(distributions, 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, iter, seed = 1, chains = 4)
```

Arguments

network	A [multinma::nma_data] object
L	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)
```

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	<i>Generate network data</i>
------------------	------------------------------

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	<i>H-function for FPs</i>
---	---------------------------

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

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

Description

Hazard Ratio

Usage

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

Arguments

<code>TTE</code>	A TTE dataframe
<code>strata</code>	A strata variable

<code>km_estimates</code>	<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

<code>TTE</code>	A TTE dataframe
------------------	-----------------

Value

A [PCNMA::km_obj] object

phi	<i>Fractional Polynomial Function</i>
-----	---------------------------------------

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	<i>Plot a fitted distributions object</i>
--------------------------	---

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 = 0.5,
  linewidth = 1,
  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 model*

Description

Plots for an NMA model

Usage

```
## S3 method for class 'fitted_model'
plot(model, type = "trace", pars = parsForStan)
```

Arguments

model	A ‘PancSurv::fitted_model‘ obj
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_aic

Plot AIC Bar plot

Description

Plot AIC Bar plot

Usage

`plot_aic(model)`

Arguments

`model` A [PCNMA::fitted_distribution] object

plot_network

Plot a network of evidence

Description

Plot a network of evidence

Usage

`plot_network(network, ...)`

Arguments

`network` A network dataset

save_nma_summary *Write model summary to xlsx*

Description

Write model summary to xlsx

Usage

```
save_nma_summary(nma, PARAMCD, effects, model_type)
```

Arguments

nma	A [mutinma::stan_data] object
PARAMCD	one of OS or PFS
model_type	one of parametric or fp
effect	type

summary.fitted_distribution *Summary of a set of fitted models*

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_distributions' object.
AIC	Returns the AIC scores for a set of models
median	Returns a table of median estimates for a set of models

summary.hr_obj	<i>Summarise a Hazard Ratio</i>
----------------	---------------------------------

Description

Summarise a Hazard Ratio

Usage

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

summary.km_obj	<i>Summarise KM data</i>
----------------	--------------------------

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_aic, 9
plot_network, 9
save_nma_summary, 10
summary.fitted_distribution, 10
summary.hr_obj, 11
summary.km_obj, 11

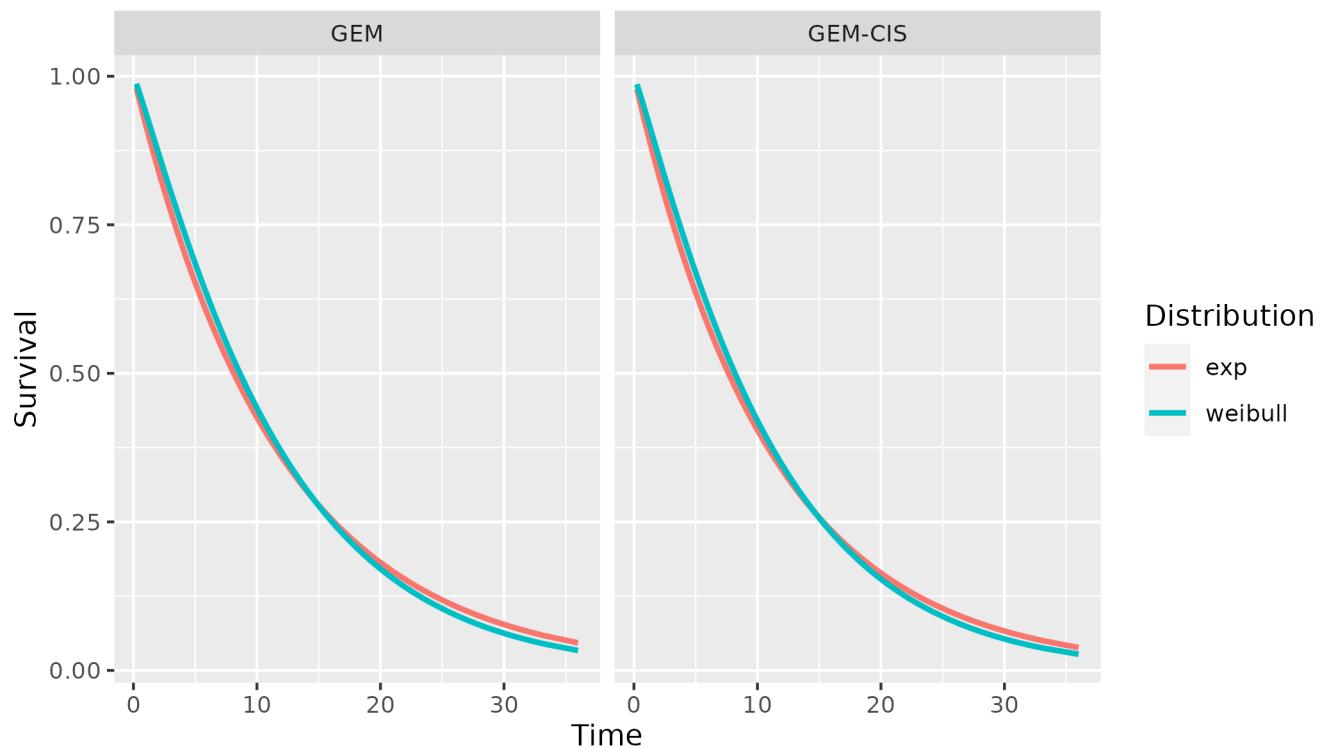
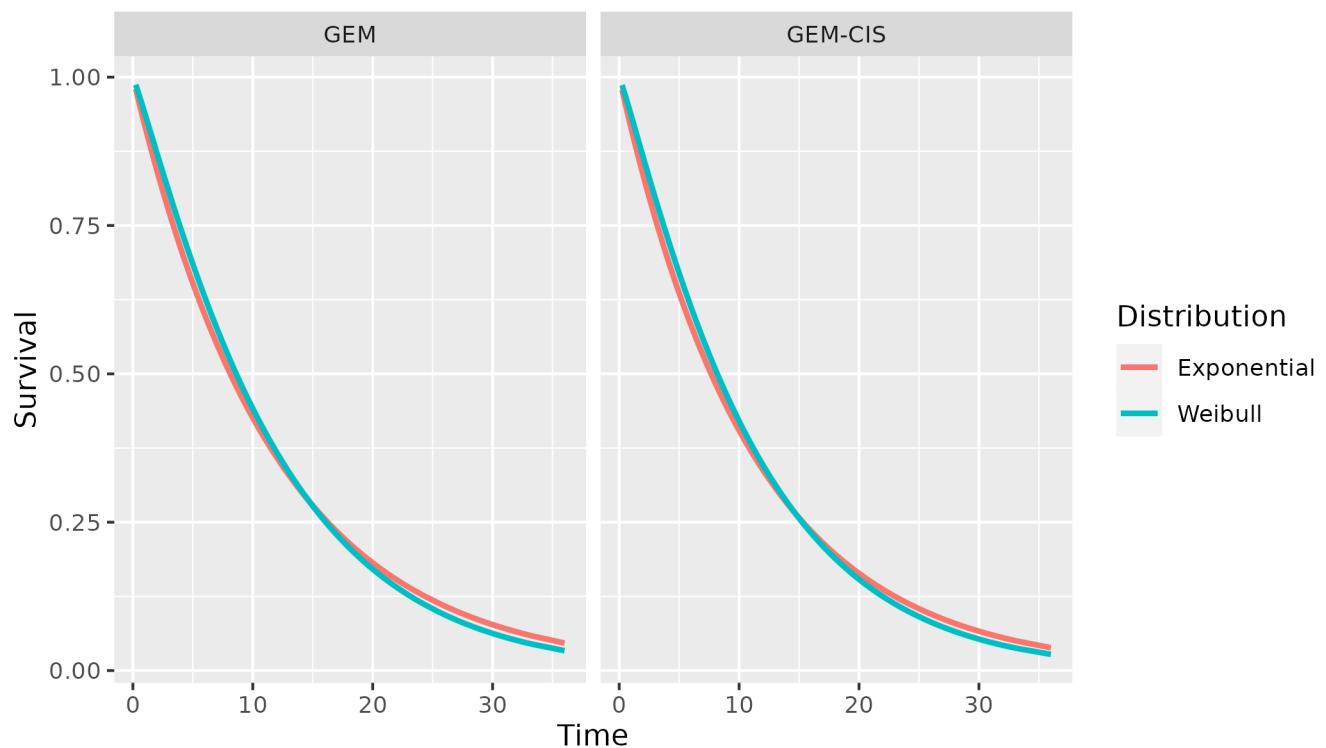


Figure C.1: Fitted Models as in Flexsurv

Figure C.2: Fitted Models with *fit_distribution*