GENERALIZED MODIFIED WEIBULL AND EXPONENTIATED WEIBULL EXPONENTIAL DISTRIBUTIONS FOR CURE FRACTION MODELS OF CANCER PATIENTS

By

MOHAMED ELAMIN ABDALLAH MOHAMED ELAMIN OMER

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfillment of the Requirements for the Degree of Doctor of Philosophy

November 2023

FS 2023 22

DEDICATION

This thesis is dedicated to the four dearest people who have meant and continue to mean a lot to me. Their memories continue to guide my life even though they are no longer here. First and foremost, I would like to thank my father, Abdallah Mohamed Elamin, who taught me the value of work ethic and whose love for me knew no limits. I appreciate you, ”dad,” and I won't ever forget you.

My mother, Asia Bit Ahmed, who raised me, loved me and helped me learn how to speak Arabic, my native language, comes next. Regrettably, despite my best efforts, we couldn't travel to Mecca together for the pilgrimage.

I also wish to commemorate my ”little” brother Hisham, who has left our loving eyes and a vacuum in our life that can never be filled. Despite the shortness of your life, I will ensure that your memory is preserved for as long as I am alive. I wish you pleasure and tranquility in paradise!

Last but not least, I dedicate this to Mohd Adam Bakri, a professor of mine. Whose life was abruptly taken from him at an early age. I will always be grateful for his excitement, never-ending encouragement, and support, all of which helped many of us succeed in our PhDs, something we had never thought was possible. There are no adequate words to express how grateful I am to him for his invaluable and persistent advice and guidance.

Beyond words, I miss every one of you. I pray that Allah (SWT) may grant you all Jannah Al-Firdaus, Ameen.

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfillment of the requirement for the degree of Doctor of Philosophy

GENERALIZED MODIFIED WEIBULL AND EXPONENTIATED WEIBULL EXPONENTIAL DISTRIBUTIONS FOR CURE FRACTION MODELS OF CANCER PATIENTS

By

MOHAMED ELAMIN ABDALLAH MOHAMED ELAMIN OMER

November 2023

Chairman: Mohd Shafie bin Mutafa, PhD
Faculty: Science

This research aims to develop a parametric cure model for lifetime data in the presence of right and interval- censored data with fixed predictors. The research begins by extending the existing Mixture Cure Model (MCM), utilizing Generalized Modified Weibull (GMW) and Exponentiated Weibull Exponential (EWE) distributions to accommodate both right- and interval-censored data with fixed covariates.

Bounded Cumulative Hazard (BCH) and the Geometric Non-Mixture Cure (GeNMC) models, are also explored, offering alternative approaches in cure modelling methodologies. These models are developed based on GMW and EWE distributions, are extended in the presence of right and interval censored data with fixed covariate.

Maximum likelihood estimation (MLE) method is employed to estimate model parameters. Simulation studies are carried out to assess the performance of the MLE estimates. The MLE performance is evaluated using bias, standard error (SE), and root mean square error (RMSE) metrics across varying sample sizes and censoring proportions. The width of the interval (len) for the interval-censored data (observational gap times) is also being considered (len=0.5). The results of the simulation studies reveal increased bias, SE, and RMSE of the estimates with higher censoring proportions and decreased sample sizes. Moreover, the MLE demonstrates efficiency, evidenced by declining RMSE values with increasing sample sizes across all censoring proportions.

To further support the findings of the simulation studies, four real-life datasets are utilized, sourced from cancer and smoking studies. The first dataset comprises of right-censored observations from a bladder cancer study. The second dataset is an interval-censored data taken from a smoking cessation study. This dataset includes smoking relapse times that were collected annually over a 5-year follow-up period from participants living in 51 zip code areas in the South Eastern region of Minnesota, USA. The third dataset includes right-censored data from a study on leukemia, focusing on treatment as the covariate. The fourth dataset is a right-censored data related to melanoma cancer, considering sex, treatment, and age as covariates.

Comparing the MCM, BCH, and GeNMC models based on GMW, EWE, Fréchet, and Gompertz distributions using bladder data, the results indicate that the MCM, BCH, and GeNMC models based on the EWE distribution performed better than the other competing models in this study. While the GMW distribution with the three cure models provides a slightly better fit than the EWE distribution, considering smoking cessation data. For leukemia data, both GMW and EWE distributions emerge as best choices for modeling the survival times of susceptible patients. For the melanoma data, while all models show similar outcomes, the MCM model with the EWE distribution exhibits the best fit.

Keywords: Parametric cure model, Right-and interval-censored data, Maximum likelihood estimation method, Mixture cure model, Non-mixture cure model, Generalized modified Weibull distribution, Exponentiated Weibull exponential distribution

SDG: GOAL 3: Good Health and Well-Being

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

TABURAN WEIBULL TERITLAK TERUBAHSUAI DAN EKSPONEN WEIBULL TEREKSPONEN BAGI MODEL PECAHAN SEMBUH PESAKIT KANSER

Oleh

MOHAMED ELAMIN ABDALLAH MOHAMED ELAMIN OMER

November 2023

Pengerusi: Mohd Shafie bin Mutafa, PhD
Fakulti: Sains

Penyelidikan ini bertujuan untuk membangunkan model penyembuh parametrik untuk data sepanjang hayat dengan kehadiran data tertapis kanan dan tertapis selang dengan peramal tetap. Penyelidikan bermula dengan melanjutkan Model Penyembuh Campuran (MCM) sedia ada, menggunakan taburan Weibull Terubahsuai Teritlak (GMW) dan taburan Eksponen Weibull Tereksponen (EWE) bagi menampung kedua-dua data tertapis selang dan tertapis kanan dengan kovariat tetap.

Model Bahaya Kumulatif Terbatas (BCH) dan Model Penyembuh Bukan Campuran Geometrik (GeNMC), juga diterokai, menawarkan pendekatan alternatif dalam metodologi pemodelan penyembuhan. Model-model ini dibangunkan berdasarkan taburan GMW dan EWE, dilanjutkan dengan kehadiran data tertapis kanan dan tertapis selang dengan kovariat tetap.

Kaedah Anggaran Kebolehjadian Maksimum (MLE) digunakan untuk menganggar parameter model. Kajian simulasi dijalankan untuk menilai prestasi penganggar MLE. Prestasi MLE dinilai menggunakan metrik pincang, ralat piawai (SE) dan punca min kuasa dua ralat (RMSE) merentas pelbagai saiz sampel dan kadaran penapisan. Lebar selang (len) bagi data tertapis selang (jurang masa cerapan) juga di pertimbangkan (len=0.5). Keputusan kajian simulasi menunjukkan bahawa pincang, SE dan RMSE meningkat dengan kadaran penapisan yang lebih tinggi dan pengurangan saiz sampel. Tambahan pula, MLE menunjukkan kecekapan, dibuktikan dengan penurunan nilai RMSE dengan peningkatan saiz sampel merentas semua kadaran penapisan.

Selanjutnya, empat set data kehidupan sebenar daripada punca kajian kanser dan merokok, digunakan untuk menyokong keputusan kajian simulasi. Set data pertama terdiri daripada cerapan tertapis kanan dari kajian kanser pundi kencing. Set data kedua adalah data tertapis selang diambil daripada kajian pemberhentian merokok. Set data ini merangkumi masa berulang merokok yang dikumpul setiap tahun sepanjang lima tahun tempoh susulan peserta yang tinggal di kawasan poskod 41 dalam Wilayah Timur Selatan Minnesota, USA. Set data ketiga termasuk data tertapis kanan daripada kajian leukimia, memfokus rawatan sebagai kovariat. Set data keempat adalah data tertapis kanan berkaitan kanser melanoma, dengan mempertimbangkan jantina, rawatan dan umur sebagai faktor kovariat.

Membandingkan model MCM, BCH dan GeNMC berdasarkan taburan GMW, EWE, Frechet dan Gompertz menggunakan data pundi kencing, keputusan menunjukkan bahawa prestasi model MCM, BCH dan GeNMC berdasarkan taburan EWE ada-lah lebih baik daripada semua model pesaing yang lain dalam kajian ini. Manakala taburan GMW dengan tiga model penyembuh memberikan penyuaian yang sedikit lebih baik daripada taburan EWE, dengan mengambil kira data pemberhentian merokok. Bagi data leukemia, kedua-dua taburan GMW dan EW muncul sebagai pilihan terbaik bagi memodelkan masa daya tahan bagi pesakit yang terdedah. Bagi data melanoma, sementara semua model menunjukkan hasil yang sama, model MCM dengan taburan EWE mempamerkan kesesuaian yang terbaik.

Kata kunci: Model penyembuh parametrik, Data tertapis kanan dan tertapis selang, Kaedah Anggaran Kebolehjadian Maksimum, Taburan Weibull Terubahsuai Teritlak, Taburan Eksponen Weibull Tereksponen, Model penyembuhan campuran, Model penyembuhan bukan campuran

SDG: MATLAMAT 3: Kesihatan dan Kesejahteraan Baik

ACKNOWLEDGEMENTS

In the name of Allah, Most Gracious, Most Merciful

First and foremost, I want to express my sincere gratitude to Allah for giving me the perseverance and fortitude necessary to finish this thesis.

Words cannot express my great appreciation to Assoc. Prof. Dr. Mohd Rizam Abu Bakar, my first supervisor, for his invaluable patience, encouragement, and feedback. In addition, I'm deeply indebted to Dr. Mohd Shafie Mustafa, who supervised me after the retirement of my first supervisor, for his guidance and generous support. Moreover, I would like to express my heartfelt gratitude to Prof. Dr. Habshah Midi, Dr. Norhaslinda Ali, and Dr. Nur Haizum Abd Rahman for their invaluable assistance and guidance as members of my supervisory committee.

Finally, I must mention my family, especially my wife, Dalia Abusin, my brothers, and sisters, for I would be remiss if I did not. Their confidence in me has kept me motivated and upbeat throughout this journey.

Last but not least, I would like to extend my sincere thanks to the staff of the Department of Mathematics and Statistics, Faculty of Science, Universiti Putra Malaysia for their help and cooperation.

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

Mohd Shafie bin Mutafa, PhD
Senior Lecturer
Faculty of Science
Universiti Putra Malaysia
(Chairperson)

Norhaslinda binti Ali, PhD
Senior Lecturer
Faculty of Science
Universiti Putra Malaysia
(Member)

Nur Haizum binti Abd Rahman, PhD
Senior Lecturer
Faculty of Science
Universiti Putra Malaysia
(Member)

ZALILAH MOHD SHARIFF, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 11 July 2024

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AFT	Accelerated Failure Time
AIC	Akaike Information Criterion
ALL	Acute Lymphocytic Leukemia
Allo-SCT	Allogeneic Stem Cell Transplantation
BCH	Bounded Cumulative Hazard
BIC	Bayesian Information Criterion
BMT	Bone Marrow Transplant
CNS	Central Nervous System
E	Exponential
ECOG	Eastern Cooperative Oncology Group
EE	Exponential Exponential
EEE	Exponentiated Exponential Exponential
EM	Expectation-Maximization
ERE	Exponentiated Rayleigh Exponential
EV	Extreme Value
EW	Exponentiated Weibull
EWE	Exponentiated Weibull Exponential
GeNMC	Geometric Non-mixture Cure
GMW	Generalized Modified Weibull
GR	Generalized Rayleigh
HQIC	Hannan-Quinn Information Criterion
HLA	Human Leukocyte Antigens
IFN	Interferon
iid	Independent and Identically Distributed
LC	Logistic/Cox

MCM	Mixture Cure Model
MGF	Moment Generating Function
MLE	Maximum Likelihood Estimation
MW	Modified Weibull
NMCM	Non-mixture cure model
PH	Proportional Hazards
SE	Standard error
R	Rayleigh
RE	Rayleigh Exponential
RMSE	Root Mean Square Error
SE	Standard Error
W	Weibull
WE	Weibull Exponential

CHAPTER 1

INTRODUCTION

1.1 Background of Study

Survival analysis is a branch of statistics that studies how long it will take for an event to occur. This field of statistics was born out of the need to track the impact of medical treatments on patient's survival in clinical trials. For instance, consider the case of a group of cancer patients who are given a new form of treatment. The results of the treatment may be analyzed in terms of the patients' life expectancy using survival analysis. Depending on the type of application, survival analysis is also known as, time-to-event analysis, lifetime data analysis, event history analysis, duration analysis, and reliability analysis.

One of the obstacles of survival analysis is that only a portion of the study group will have experienced the event by the end of the follow-up period. Therefore, survival durations will be unknown for a subset of the study group. This phenomenon is known as censoring, which arises when a study participant has not yet encountered a relevant event (such as relapse or death) by the end of the study. Censoring may occur for various reasons. For instance, the study participant fails to follow-up, withdraws from the study, or is still alive or disease-free at the end of the study period.

Survival models are commonly used in medical research to evaluate time-to-event data, in which individuals are tracked over some time, and the time until an event of interest occurs is recorded. For instance, a study to model the time after colon cancer patients' treatment until death or the time from the first heart attack to the second. It is usually believed that if a study subject is followed long enough, he or she would eventually experience the event of interest. Nevertheless, even after a lengthy period of time, the event may not occur in some subjects. For example, in a breast or prostate cancer study, it is typical for some patients to never have cancer relapse following treatment. In this situation, the patients are not censored in the usual sense and are thus firmly believed to be cured. Therefore, classical survival models such as the accelerated failure time and Cox proportional hazard models are inappropriate for such instances and this type of data. As a result, cure rate models have been mostly developed to deal with this kind of data. In the cure model formulation, the censored observation group is split into two subgroups: those who are event-free, and therefore, cured or immune; and those who will ultimately experience events if they are observed for a sufficiently long time.

In the literature, there are two main types of cure models that have been suggested to fit lifetime data in medical studies, namely, the mixture cure model introduced by

Farewell (1977) and Farewell (1982), and the promotion time cure model suggested by Tsodikov et al. (1996). These models can be used to analyze real-life data in different domains other than medicine, such as economics, reliability, criminology, sociology education, and marketing, to name a few. The modeling approach varies based on the researcher's event of interest; the common idea is to observe time until the event, but for some subjects, the event will never occur.

At the end of the observation period, the survival times of some individuals may be censored. As a result, if the follow-up is long enough, the dataset might contain cured individuals. However, broadly, cured individuals include a subset of censored individuals. The challenge in fitting cure models on survival data is that the existence of a cure rate in the sample is not evident. Therefore, it is recommended to test whether there is an immune portion in the given dataset before fitting the cure models, as well as if the follow-up period is long enough (Maller and Zhou, 1996).

1.2 Scope of Study

This thesis focuses on the problems of estimating the cure fraction and the influence of covariates on the probability of being cured and the survival times of susceptible subjects. This study is organized into two parts. The first part will be devoted to extending the parametric cure models to incorporate right and interval-censored data with or without fixed covariates. The analysis is based on two parametric distributions, recently introduced in the literature: Generalized Modified Weibull and Exponentiated Weibull exponential distributions. Comprehensive simulation studies were carried out to determine the parameter estimates of these models, and this involved maximum likelihood estimation techniques, assessing the bias, standard error, and root mean square error of the parameter estimates across various sample sizes and levels of censoring. In the other part, the practical application of the newly proposed models is illustrated by using four datasets from oncology and clinical trial studies.

1.3 Problem Statement

Many cure models have been developed to analyse survival data with long-term survivors. The parametric technique is one method for predicting the cure probability and survival function for uncured subjects. Numerous parametric distributions, including Weibull, Fréchet, and lognormal, have been applied in both Mixture Cure Model (MCM) and Non-Mixture Cure model, such as Bounded Cumulative Hazard (BCH) and Geometric Non-Mixture Cure (GeNMC) model , for analysing survival data involving cure proportion. These distributions are preferred for their flexibility in hazard functions and the ease of parameter estimation. However, their utility comes with some limitations:

- They may encounter challenges in accurately modelling complex hazard functions, and their ability to capture a diverse range of shapes in survival functions may be constrained, rendering them less suitable for scenarios with intricate hazard patterns.

- These distributions are not well-suited for situations where hazards deviate from a strictly proportional pattern, leading to inadequate fits when hazard rates undergo non-proportional changes over time.

- Their effectiveness diminishes when handling heterogeneous datasets characterized by varying failure behaviors across subsets. In such cases, these distributions struggle to accommodate the diverse patterns present in the data, limiting their efficacy in capturing the inherent complexity of heterogeneous datasets.

Many statistical estimation methods were used so far in the frame of cure models estimation, such as the Maximum Likelihood Estimation (MLE), Bayesian estimation, Expectation-Maximization algorithm, and Profile Likelihood. These methods offer distinct approaches to parameter estimation for cure models. Collectively, they address diverse statistical scenarios, enhancing the robustness and flexibility of parameter estimation across various fields.

In this research, we will consider two recent parametric distributions, namely the Generalized Modified Weibull (GMW) distribution and the Exponentiated Weibull Exponential (EWE) distribution. It is noteworthy that the GMW distribution has not been employed in the framework of the GeNMC model, and similarly, the EWE distribution has not been used yet with the MCM, BCH model, and the GeNMC model. Compare to Weibull, Fréchet, and lognormal distributions, the GMW and EWE distributions exhibit the following characteristics:

- Their extra parameters, enhancing the ability to model intricate hazard functions. This increased flexibility enables these distributions to effectively capture a broader spectrum of shapes in the survival function, making them applicable in scenarios where the hazard cannot be adequately described by Weibull, Fréchet, and lognormal distributions.

- While the Weibull, Fréchet, and lognormal distributions satisfy proportional hazards property, GMW and EWE distributions offer additional flexibility, making them more suitable for situations where hazards do not strictly adhere to a proportional pattern. This adaptability proves beneficial when dealing with data exhibiting non-proportional hazards, where hazard rates experience changes as time progresses.

- GMW and EWE distributions excel in handling heterogeneous datasets characterized by varying failure behaviours across different subsets. Their inherent flexibility allows them to accommodate the diverse patterns present in such

datasets, making them well-suited for capturing the complexity inherent in heterogeneous data, which is not the same in case for the Weibull, Fréchet, and lognormal distributions.

Maximum Likelihood Estimation (MLE) is considered in this study due to its numerous advantages. Firstly, MLE provides efficient estimates, meaning that under specific conditions, the estimates derived from MLE are asymptotically unbiased and have the smallest variance compared to all other unbiased estimators (Ren and Wang, 2023). Secondly, MLE estimators exhibit desirable asymptotic properties such as consistency, asymptotic normality, and asymptotic efficiency, indicating that as the sample size increases, MLE estimates converge towards the true parameter values (Kim et al., 2010). Additionally, MLE demonstrates precision when dealing with continuous data following a uniform distribution (Mindrila, 2010). Furthermore, MLE offers flexibility as it does not require distributional assumptions beyond those implied by the likelihood function, setting it apart from methods reliant on specific distributional assumptions (Chan, 2021). MLE plays a crucial role across various statistical modeling methods, particularly in nonlinear modeling with non-normally distributed data (Myung, 2003). Moreover, MLE methods possess favorable mathematical properties and optimality characteristics, evolving into minimum variance unbiased estimators as sample size increases (Schneider, 2018). They also exhibit approximate normal distributions and sample variances, enabling the construction of confidence intervals (Heckert et al., 2002). Finally, under specific conditions, MLE achieves the Cramér-Rao lower bound, making it the most efficient estimator among all unbiased estimators, ensuring optimal performance in parameter estimation tasks (Pfanzagl, 2011). These advantages collectively justify the consideration of MLE in this study.

The expected novelty of the findings of this research is outlined as follows:

1. The utilization of the GMW and EWE distributions in the context of MCM, BCH and GeNMC models for cancer patients is anticipated to enhance the statistical rigor of cancer research. By employing advanced parametric modeling techniques, this research seeks to improve the precision and reliability of estimates related to cure rates, offering a more nuanced understanding of the dynamics of cancer patient outcomes.

2. This research is expected to contribute to the advancement of cure fractions models for cancer patients. By extending MCM, BCH and GeNMC models with the GMW and EWE distributions to accommodate right and interval-censored data in the presence of fixed covariates, we aim to enhance the current understanding of the factors influencing cure rates and survival patterns in cancer populations. This novel application could reveal insights into the underlying mechanisms of long-term survival and remission.

3. The comparative analysis of the GMW and EWE distributions within the MCM, BCH and GeNMC models framework is anticipated to add a unique

dimension to the study. By assessing the performance of these distributions in modeling cure fractions, the research aims to identify which distribution better aligns with the cancer datasets; Bone Marrow Transplant data (BMT) and Melanoma E1684 data. This comparative aspect contributes to the methodological discourse surrounding the choice of distributions in cure models.

4. The findings from this study are expected to provide valuable insights for cancer treatment strategies. Understanding the distributional characteristics associated with cure fractions can inform medical professionals and researchers about the efficacy of existing treatments and the potential for developing new therapeutic approaches tailored to specific patient subgroups.

1.4 Research Objectives

- To extend the Mixture Cure Model (MCM), Bounded Cumulative Hazard (BCH) model, and Geometric Non-Mixture Cure (GeNMC) model with the Generalized Modified Weibull (GMW) and Exponentiated Weibull Exponential (EWE) distributions, accommodating right and interval-censored data in the presence of fixed covariates.
- To apply the Maximum Likelihood Estimation (MLE) method for estimating the parameters of the MCM, BCH, and GeNMC models based on the GMW and EWE distributions.
- To evaluate the effectiveness of the proposed models through simulation studies, examining the bias, standard error, and root mean square error of parameter estimates across various sample sizes and levels of censoring.
- To illustrate the application of the proposed models on two medical datasets, namely the BMT and Melanoma cancer datasets, in order to assess the efficacy of the models using real-life data.

1.5 Outline of the Thesis

There are seven chapters in this thesis. A review of literature relevant to the current work is presented in Chapter 2. To trace the development of parametric cure models, a brief review of these models is presented in this chapter. Special focus in this literature review is given to the research conducted on estimation methods, particularly examining approaches that accommodate fixed covariates in the presence of right and interval censored data.

In Chapter 3, we extend the existing mixture cure model to incorporate right- and

interaval-censored data with fixed covariates. The survival times of susceptible individuals are modeled using the generalized modified Weibull (GMW) and exponentiated Weibull exponential (EWE) distributions. The performance of this model is compared at different sample sizes and various censoring rates via extensive simulation studies.

Chapter 4 introduces a bounded cumulative hazard model (BCH) with fixed covariates based on the GMW and EWE distributions with different types of censored observations. Also, this chapter introduces a modified class of cure rates models that can be utilized as an alternative to the MCM and BCH models. Furthermore, this chapter evaluates the performance of the MCM, BCH, and GeNMC models across different distributions, employing multiple criteria to determine their fit. This comparison utilizes two real-life datasets: bladder cancer data and smoking cessation data.

Chapters 5 and 6 provide an in-depth exploration of the practical applications of parametric cure models in oncology research and clinical trials. These chapters investigate the MCM, BCH, and GeNMC models, considering different parametric distributions, censored observations, and fixed covariates. The analysis aims to understand the influence of predictors on the probability of being cured and their impact on the survival times of susceptible individuals. In Chapter 5, the focus is on the application of parametric mixture cure models introduced in Chapter 3, specifically using the Bone Marrow Transplant (BMT) study dataset. Chapter 6, on the other hand, concentrates on the implementation of parametric non-mixture cure models discussed in Chapter 4, utilizing the E1684 Melanoma dataset from a phase III clinical trial conducted by the Eastern Cooperative Oncology Group (ECOG). These chapters provide valuable insights and practical guidance for researchers and clinicians in the field of oncology, enabling them to apply these parametric cure models effectively in their own studies and trials.

Chapter 7 concludes this thesis by summarizing the key findings and offering some recommendations for future studies.

The road map of this research is shown in Figure 1.1.

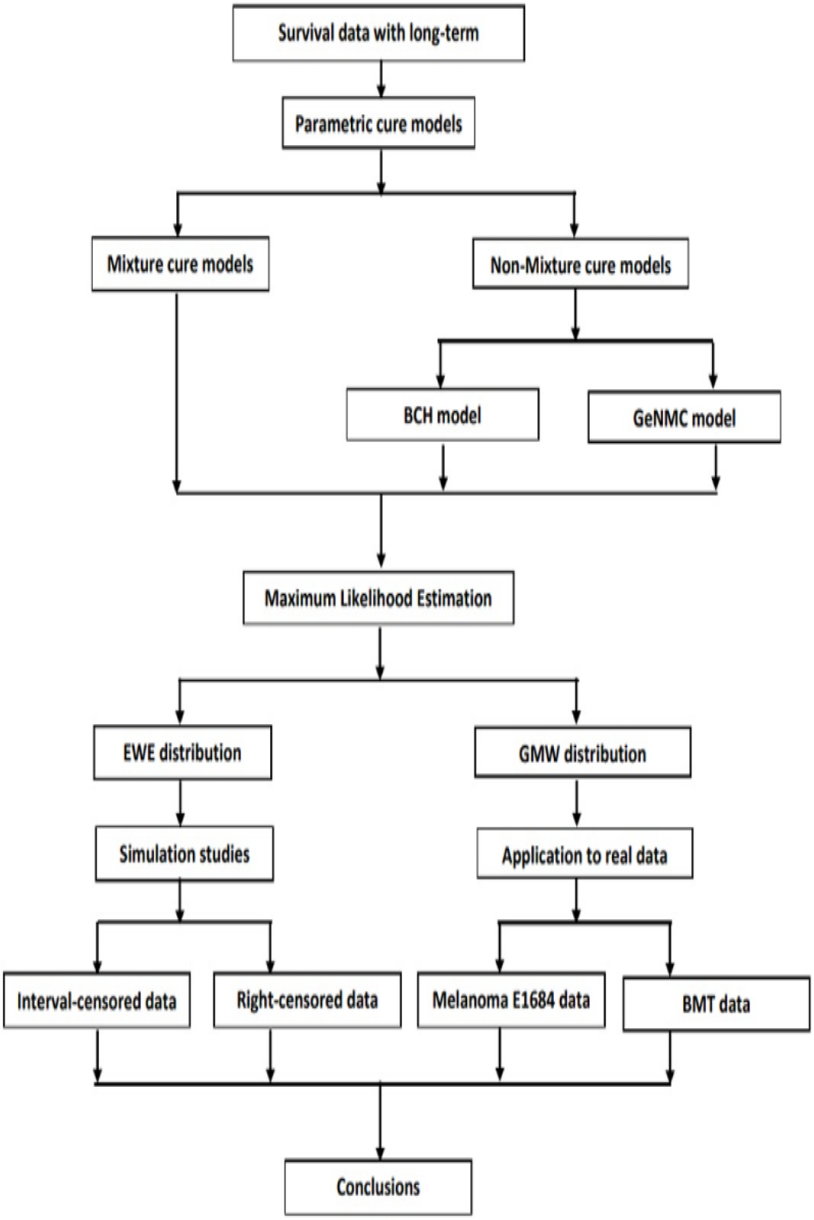

Figure 1.1: Conceptual road map

REFERENCES

Abreu, A. M. and Rocha, C. S. (2013). A parametric cure model with covariates. In *Advances in Regression, Survival Analysis, Extreme Values, Markov Processes and Other Statistical Applications*, pages 37–45. Springer.

Ahmed, A. O. M. (2021). Bayesian estimations of exponential distribution based on interval-censored data with a cure fraction. *Journal of Mathematics*, 2021:1–11.

Aho, K., Derryberry, D., and Peterson, T. (2017). A graphical framework for model selection criteria and significance tests: refutation, confirmation and ecology. *Methods in Ecology and Evolution*, 8(1):47–56.

Aljawadi, B. A. I., Bakar, M. R. A., Ibrahim, N. A., and Midi, H. (2011). Parametric estimation of the cure fraction based on bch model using left-censored data with covariates. *Modern Applied Science*, 5(3):103.

Almalki, S. J. and Nadarajah, S. (2015). Comparing the exponentiated and generalized modified weibull distributions. *Journal of Data Science*, 13(4):713–731.

Alzaghal, A., Lee, C., and Famoye, F. (2013). Exponentiated weibull-exponential distribution: Properties and applications. *Journal of Applied Statistical Science*, 21(1):113.

American Cancer Society (Retrieved 12/03/2023). Cancer facts and figures 2023. https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2023/2023-cancer-facts-and-figures.pdf.

Amico, M. and Van Keilegom, I. (2018). Cure models in survival analysis. *Annual Review of Statistics and Its Application*, 5:311–342.

Anderson, T. W. and Darling, D. A. (1954). A test of goodness of fit. *Journal of the American statistical association*, 49(268):765–769.

Arasan, J. and Adam, M. B. (2014). Double bootstrap confidence interval estimates with censored and truncated data. *Journal of Modern Applied Statistical Methods*, 13(2):22.

Arasan, J. and Midi, H. (2023). Bootstrap based diagnostics for survival regression model with interval and right-censored data. *Austrian Journal of Statistics*, 52(2):66–85.

Baak, J., Li, H., and Guo, H. (2022). Clinical and biological interpretation of survival curves of cancer patients, exemplified with stage iv non-small cell lung cancers with long follow-up. *Frontiers in Oncology*, 12:837419.

Bailey, C., Panfil, K., and Kirkpatrick, K. (2021). Hazard function effects on promoting self-control in variable interval time-based interventions in rats. *Journal of the experimental analysis of behavior*, 116(3):279–299.

Bakar, M. R. A., Daud, I., Ibrahim, N. A., and Rahmatina, D. (2006). Estimating a logistic weibull mixture models with long-term survivors. *Jurnal Teknologi*, pages 57–66.

Balogun, O. S., Olaleye, S. A., Gao, X.-Z., and Toivanen, P. (2020). A novel generalized form of cure rate model for an infectious disease with co-infection. In *International Conference on Intelligent Systems Design and Applications*, pages 69–79. Springer.

Barriga, G. D., Cancho, V. G., Garibay, D. V., Cordeiro, G. M., and Ortega, E. M. (2019). A new survival model with surviving fraction: An application to colorectal cancer data. *Statistical methods in medical research*, 28(9):2665–2680.

Bartolucci, F., Scrucca, L., et al. (2010). Point estimation methods with applications to item response theory models. In *International Encyclopedia of Education. 3rd edition*, pages 366–373. Oxford: Elsevier.

Beran, R. (1981). Efficient robust estimates in parametric models. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 55:91–108.

Berkson, J. and Gage, R. P. (1952). Survival curve for cancer patients following treatment. *Journal of the American Statistical Association*, 47(259):501–515.

Beyhum, J., El Ghouch, A., Portier, F., and Van Keilegom, I. (2022). On an extension of the promotion time cure model. *The Annals of Statistics*, 50(1):537–559.

Boag, J. W. (1949). Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *Journal of the Royal Statistical Society. Series B (Methodological)*, 11(1):15–53.

Borges, P., Rodrigues, J., and Balakrishnan, N. (2012). Correlated destructive generalized power series cure rate models and associated inference with an application to a cutaneous melanoma data. *Computational Statistics & Data Analysis*, 56(6):1703–1713.

Bremhorst, V. and Lambert, P. (2016). Flexible estimation in cure survival models using bayesian p-splines. *Computational Statistics & Data Analysis*, 93:270–284.

Brownlee, J. (2019). A gentle introduction to expectation-maximization (em algorithm). *Machine Learning Mastery, Oct*, 31.

Brusa, L., Bartolucci, F., and Pennoni, F. (2023). Tempered expectation-maximization algorithm for the estimation of discrete latent variable models. *Computational Statistics*, 38(3):1391–1424.

Cai, C., Zou, Y., Peng, Y., and Zhang, J. (2012). smcure: An r-package for estimating semiparametric mixture cure models. *Computer methods and programs in biomedicine*, 108(3):1255–1260.

Cancho, V. G., Bedia, E. C., Cordeiro, G. M., Prataviera, F., Ortega, E. M., and Santo, A. P. (2023). A survival regression with cure fraction applied to cervical cancer. *Computational Statistics*, 38(1):403–418.

Cancho, V. G., de Castro, M., and Dey, D. K. (2013). Long-term survival models with latent activation under a flexible family of distributions.

Cancho, V. G., Louzada, F., and Barriga, G. D. (2012). The geometric birnbaum–saunders regression model with cure rate. *Journal of Statistical Planning and Inference*, 142(4):993–1000.

Carrasco, J. M., Ortega, E. M., and Cordeiro, G. M. (2008). A generalized modified weibull distribution for lifetime modeling. *Computational Statistics & Data Analysis*, 53(2):450–462.

Chan, S. H. (2021). *Introduction to probability for data science*. Michigan Publishing.

Chen, M.-H., Ibrahim, J. G., and Sinha, D. (1999). A new bayesian model for survival data with a surviving fraction. *Journal of the American Statistical Association*, 94(447):909–919.

Clark, T. G., Bradburn, M. J., Love, S. B., and Altman, D. G. (2003). Survival analysis part i: basic concepts and first analyses. *British journal of cancer*, 89(2):232–238.

Clyde, M., Berger, J., Bullard, F., Ford, E., Jefferys, W., Luo, R., Paulo, R., and Loredo, T. (2007). Current challenges in bayesian model choice. In *Statistical challenges in modern astronomy IV*, volume 371, page 224.

Cook, T. D., Steiner, P. M., and Pohl, S. (2009). How bias reduction is affected by covariate choice, unreliability, and mode of data analysis: Results from two types of within-study comparisons. *Multivariate Behavioral Research*, 44(6):828–847.

Cooner, F., Banerjee, S., Carlin, B. P., and Sinha, D. (2007). Flexible cure rate modeling under latent activation schemes. *Journal of the American Statistical Association*, 102(478):560–572.

Cooner, F., Banerjee, S., and McBean, A. M. (2006). Modelling geographically referenced survival data with a cure fraction. *Statistical methods in medical research*, 15(4):307–324.

Corbiere, F., Commenges, D., Taylor, J. M., and Joly, P. (2009). A penalized likelihood approach for mixture cure models. *Statistics in medicine*, 28(3):510–524.

De la Cruz, R., Fuentes, C., and Padilla, O. (2022). A bayesian mixture cure rate model for estimating short-term and long-term recidivism. *Entropy*, 25(1):56.

De Pascoa, M. A., Ortega, E. M., and Cordeiro, G. M. (2011). The kumaraswamy generalized gamma distribution with application in survival analysis. *Statistical methodology*, 8(5):411–433.

Demicheli, R., Hrushesky, W., Retsky, M., and Biganzoli, E. (2020). Interpreting breast cancer survival data by the hazard function: remarkable findings from event dynamics. *Medicina*, 56(9):468.

Diop, A., Diop, A., and Dupuy, J.-F. (2011). Maximum likelihood estimation in the logistic regression model with a cure fraction.

Dudley, W. N., Wickham, R., and Coombs, N. (2016). An introduction to survival statistics: Kaplan-meier analysis. *Journal of the advanced practitioner in oncology*, 7(1):91.

Dwivedi, N., Sachdeva, S., and Sulania, A. (2016). Bathtub concept in health sciences: A comment. *Indian Journal of Health Sciences and Biomedical Research kleu*, 9(1):117–120.

Elgarhy, M., Shakil, M., and Kibria, G. (2017). Exponentiated weibull-exponential distribution with applications. *Applications and Applied Mathematics: An International Journal (AAM)*, 12(2):5.

Fang, H.-b., Li, G., and Sun, J. (2005). Maximum likelihood estimation in a semiparametric logistic/proportional-hazards mixture model. *Scandinavian Journal of Statistics*, 32(1):59–75.

Farewell, V. T. (1977). A model for a binary variable with time-censored observations. *Biometrika*, 64(1):43–46.

Farewell, V. T. (1982). The use of mixture models for the analysis of survival data with long-term survivors. *Biometrics*, pages 1041–1046.

Feller, W. (1991). *An introduction to probability theory and its applications, Volume 2*, volume 81. John Wiley & Sons.

Fogli, S., Arena, C., Carpi, S., Polini, B., Bertini, S., Digiacomo, M., Gado, F., Saba, A., Saccomanni, G., Breschi, M. C., et al. (2016). Cytotoxic activity of oleocanthal isolated from virgin olive oil on human melanoma cells. *Nutrition and cancer*, 68(5):873–877.

Gallardo, D. I., de Castro, M., and Gómez, H. W. (2021). An alternative promotion time cure model with overdispersed number of competing causes: An application to melanoma data. *Mathematics*, 9(15):1815.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (1995). *Bayesian data analysis*. Chapman and Hall/CRC.

Gómez, Y. M., Gallardo, D. I., Bourguignon, M., Bertolli, E., and Calsavara, V. F. (2023). A general class of promotion time cure rate models with a new biological interpretation. *Lifetime Data Analysis*, 29(1):66–86.

Greenwood, M. (1938). The first life table. *Notes and Records of the Royal Society of London*, 1(2):70–72.

Gu, Y., Sinha, D., and Banerjee, S. (2011). Analysis of cure rate survival data under proportional odds model. *Lifetime data analysis*, 17:123–134.

Gustafsson, O., Villani, M., and Stockhammar, P. (2020). Bayesian optimization of hyperparameters when the marginal likelihood is estimated by mcmc.

Hanin, L., Tsodikov, A., and Yakovlev, A. Y. (2001). Optimal schedules of cancer surveillance and tumor size at detection. *Mathematical and Computer Modelling*, 33(12-13):1419–1430.

Hassan, A. S. and Elgarhy, M. (2016). A new family of exponentiated weibull-generated distributions. *International Journal of Mathematics and its Applications*, 4(1-D):135–148.

Haybittle, J. (1959). The estimation of the proportion of patients cured after treatment for cancer of the breast. *The British journal of radiology*, 32(383):725–733.

Haybittle, J. (1965). A two-parameter model for the survival curve of treated cancer patients. *Journal of the American Statistical Association*, 60(309):16–26.

Heckert, N. A., Filliben, J. J., Croarkin, C. M., Hembree, B., Guthrie, W. F., Tobias, P., and Prinz, J. (2002). Handbook 151: Nist/sematech e-handbook of statistical methods.

Hess, K. R. and Levin, V. A. (2014). Getting more out of survival data by using the hazard function. *Clinical Cancer Research*, 20(6):1404–1409.

Hougaard, P. and Hougaard, P. (2000). *Analysis of multivariate survival data*, volume 564. Springer.

Huang, J. and Wellner, J. A. (1997). Interval censored survival data: a review of recent progress. In *Proceedings of the first Seattle symposium in biostatistics: survival analysis*, pages 123–169. Springer.

Hubbell, E., Clarke, C. A., Smedby, K. E., Adami, H.-O., and Chang, E. T. (2024). Potential for cure by stage across the cancer spectrum in the united states. *Cancer Epidemiology, Biomarkers & Prevention*, 33(2):206–214.

Ibrahim, J. G., Chen, M.-H., and Sinha, D. (2001). Bayesian semiparametric models for survival data with a cure fraction. *Biometrics*, 57(2):383–388.

Jäntschi, L. and Bolboacă, S. D. (2018). Computation of probability associated with anderson–darling statistic. *Mathematics*, 6(6):88.

Karamoozian, A., Baneshi, M. R., and Bahrampour, A. (2021). Bayesian mixture cure rate frailty models with an application to gastric cancer data. *Statistical Methods in Medical Research*, 30(3):731–746.

Karlis, D. and Xekalaki, E. (2003). Choosing initial values for the em algorithm for finite mixtures. *Computational Statistics & Data Analysis*, 41(3-4):577–590.

Kersey, J. H., Weisdorf, D., Nesbit, M. E., LeBien, T. W., Woods, W. G., McGlave, P. B., Kim, T., Vallera, D. A., Goldman, A. I., Bostrom, B., et al. (1987). Comparison of autologous and allogeneic bone marrow transplantation for treatment of high-risk refractory acute lymphoblastic leukemia. *New England Journal of Medicine*, 317(8):461–467.

Kim, J. and DeBerardinis, R. J. (2019). Mechanisms and implications of metabolic heterogeneity in cancer. *Cell metabolism*, 30(3):434–446.

Kim, Y., Kim, B., and Jang, W. (2010). Asymptotic properties of the maximum likelihood estimator for the proportional hazards model with doubly censored data. *Journal of Multivariate Analysis*, 101(6):1339–1351.

Kim, Y.-J. and Jhun, M. (2008). Cure rate model with interval censored data. *Statistics in medicine*, 27(1):3–14.

Kirkwood, J. M., Strawderman, M. H., Ernstoff, M. S., Smith, T. J., Borden, E. C., and Blum, R. H. (1996). Interferon alfa-2b adjuvant therapy of high-risk resected cutaneous melanoma: the eastern cooperative oncology group trial est 1684. *Journal of clinical oncology*, 14(1):7–17.

Klein, J. P. and Moeschberger, M. L. (2003). *Survival analysis: techniques for censored and truncated data*, volume 1230. Springer.

Klein, J. P., Van Houwelingen, H. C., Ibrahim, J. G., and Scheike, T. H. (2014). *Handbook of survival analysis*. CRC Press Boca Raton, FL:.

Kleinbaum, D. G. and Klein, M. (1996). *Survival analysis a self-learning text*. Springer.

Kuk, A. Y. and Chen, C.-H. (1992). A mixture model combining logistic regression with proportional hazards regression. *Biometrika*, 79(3):531–541.

Kutal, D. H. and Qian, L. (2018). A non-mixture cure model for right-censored data with fréchet distribution. *Stats*, 1(1):176–188.

Lai, C., Xie, M., and Murthy, D. (2003). A modified weibull distribution. *IEEE Transactions on reliability*, 52(1):33–37.

Lambert, P. and Bremhorst, V. (2019). Estimation and identification issues in the promotion time cure model when the same covariates influence long-and short-term survival. *Biometrical Journal*, 61(2):275–289.

Lambert, P. C. (2007). Modeling of the cure fraction in survival studies. *The Stata Journal*, 7(3):351–375.

Lawless, J. F. (2011). *Statistical models and methods for lifetime data*. John Wiley & Sons.

Lázaro, E., Armero, C., and Gómez-Rubio, V. (2020). Approximate bayesian inference for mixture cure models. *TEST*, 29(3):750–767.

Leão, J., Bourguignon, M., Gallardo, D. I., Rocha, R., and Tomazella, V. (2020). A new cure rate model with flexible competing causes with applications to melanoma and transplantation data. *Statistics in Medicine*, 39(24):3272–3284.

Lee, E. T. and Wang, J. (2003). *Statistical methods for survival data analysis*, volume 476. John Wiley & Sons.

Lima, C. M., Tomazella, V. L., Evangelista, A. F., Campelo, J. E., and Junior, S. C. (2023). Gamma-gompertz mixture model with cure fraction to analyze data on anglo-nubian goats with positive epg. *Small Ruminant Research*, 218:106879.

Lin, L.-H. and Huang, L.-S. (2019). Yakovlev promotion time cure model with local polynomial estimation. *arXiv preprint arXiv:1911.00741*.

Liu, X. and Xiang, L. (2021). Generalized accelerated hazards mixture cure models with interval-censored data. *Computational Statistics & Data Analysis*, 161:107248.

Looha, M. A., Zarean, E., Masaebi, F., Pourhoseingholi, M. A., and Zali, M. R. (2021). Assessment of prognostic factors in long-term survival of male and female patients with colorectal cancer using non-mixture cure model based on the weibull distribution. *Surgical Oncology*, 38:101562.

López-Cheda, A., Cao, R., Jácome, M. A., and Van Keilegom, I. (2017). Nonparametric incidence estimation and bootstrap bandwidth selection in mixture cure models. *Computational Statistics & Data Analysis*, 105:144–165.

Lu, W. (2008). Maximum likelihood estimation in the proportional hazards cure model. *Annals of the Institute of Statistical Mathematics*, 60:545–574.

Lu, W. (2010). Efficient estimation for an accelerated failure time model with a cure fraction. *Statistica Sinica*, 20:661.

Maller, R. A. and Zhou, X. (1996). *Survival analysis with long-term survivors*, volume 525. Wiley New York.

Marinho, A. R. and Loschi, R. H. (2020). Bayesian cure fraction models with measurement error in the scale mixture of normal distribution. *Statistical Methods in Medical Research*, 29(9):2411–2444.

Martinez, E. Z. and Achcar, J. A. (2018). A new straightforward defective distribution for survival analysis in the presence of a cure fraction. *Journal of Statistical Theory and Practice*, 12(4):688–703.

Martinez, E. Z., Achcar, J. A., Jácome, A. A., and Santos, J. S. (2013). Mixture and non-mixture cure fraction models based on the generalized modified weibull distribution with an application to gastric cancer data. *Computer methods and programs in biomedicine*, 112(3):343–355.

Martinez, E. Z., de Freitas, B. C. L., Achcar, J. A., Aragon, D. C., and de Oliveira Peres, M. V. (2022). Exponentiated weibull models applied to medical data in presence of right-censoring, cure fraction and covariates. *Statistics, Optimization & Information Computing*, 10(2):548–571.

McElreath, R. (2018). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Chapman and Hall/CRC.

Miller, K. D., Nogueira, L., Devasia, T., Mariotto, A. B., Yabroff, K. R., Jemal, A., Kramer, J., and Siegel, R. L. (2022). Cancer treatment and survivorship statistics, 2022. *CA: a cancer journal for clinicians*, 72(5):409–436.

Mindrila, D. (2010). Maximum likelihood (ml) and diagonally weighted least squares (dwls) estimation procedures: A comparison of estimation bias with ordinal and multivariate non-normal data. *International Journal of Digital Society*, 1(1):60–66.

Molina, K. C., Calsavara, V. F., Tomazella, V. D., and Milani, E. A. (2021). Survival models induced by zero-modified power series discrete frailty: Application with a melanoma data set. *Statistical Methods in Medical Research*, 30(8):1874–1889.

Moore, D. F. (2016). *Applied survival analysis using R*, volume 473. Springer.

Müller, U. K. (2013). Risk of bayesian inference in misspecified models, and the sandwich covariance matrix. *Econometrica*, 81(5):1805–1849.

Murray, R. P., Anthonisen, N. R., Connett, J. E., Wise, R. A., Lindgren, P. G., Greene, P. G., Nides, M. A., Group, L. H. S. R., et al. (1998). Effects of multiple attempts to quit smoking and relapses to smoking on pulmonary function. *Journal of clinical epidemiology*, 51(12):1317–1326.

Musta, E., Patilea, V., and Van Keilegom, I. (2020). A presmoothing approach for estimation in semiparametric mixture cure models. *arXiv preprint arXiv:2008.05338*.

Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of mathematical Psychology*, 47(1):90–100.

Naseri, P., Baghestani, A. R., Momenyan, N., and Akbari, M. E. (2018). Application of a mixture cure fraction model based on the generalized modified weibull distribution for analyzing survival of patients with breast cancer. *International Journal of Cancer Management*, 11(5).

National Cancer Institute (Retrieved 12/03/2023). Acute lymphocytic leukemia-cancer stat facts. https://seer.cancer.gov/statfacts/html/alyl.html.

Othus, M., Li, Y., and Tiwari, R. C. (2009). A class of semiparametric mixture cure survival models with dependent censoring. *Journal of the American Statistical Association*, 104(487):1241–1250.

Pal, S. and Aselisewine, W. (2023). A semiparametric promotion time cure model with support vector machine. *The Annals of Applied Statistics*, 17(3):2680–2699.

Pal, S., Barui, S., Davies, K., and Mishra, N. (2022). A stochastic version of the em algorithm for mixture cure model with exponentiated weibull family of lifetimes. *Journal of Statistical Theory and Practice*, 16(3):48.

Patilea, V. and Van Keilegom, I. (2020). A general approach for cure models in survival analysis.

Pedrosa-Laza, M., López-Cheda, A., and Cao, R. (2022). Cure models to estimate time until hospitalization due to covid-19: A case study in galicia (nw spain). *Applied Intelligence*, 52(1):794–807.

Peña-Ramírez, F. A., Guerra, R. R., Cordeiro, G. M., and Marinho, P. R. (2018). The exponentiated power generalized weibull: Properties and applications. *Anais da Academia Brasileira de Ciências*, 90:2553–2577.

Peng, Y. and Dear, K. B. (2000). A nonparametric mixture model for cure rate estimation. *Biometrics*, 56(1):237–243.

Peng, Y., Dear, K. B., and Carriere, K. (2001). Testing for the presence of cured patients: a simulation study. *Statistics in medicine*, 20(12):1783–1796.

Peng, Y. and Taylor, J. M. (2014). Cure models. *Handbook of survival analysis*, 34:113–134.

Peng, Y. and Xu, J. (2012). An extended cure model and model selection. *Lifetime data analysis*, 18:215–233.

Peng, Y. and Yu, B. (2021). *Cure models: methods, applications, and implementation*. Chapman and Hall/CRC.

Pfanzagl, J. (2011). *Parametric statistical theory*. Walter de Gruyter.

Rahimzadeh, M. and Kavehie, B. (2016). Promotion time cure model with generalized poisson-inverse gaussian distribution. *Journal of Biostatistics and Epidemiology*, 2(2):68–75.

Ramos, P. L., Guzman, D. C., Mota, A. L., Rodrigues, F. A., and Louzada, F. (2020). Sampling with censored data: a practical guide. *arXiv preprint arXiv:2011.08417*.

Razali, N. M., Wah, Y. B., et al. (2011). Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. *Journal of statistical modeling and analytics*, 2(1):21–33.

Ren, J. and Wang, H. (2023). *Mathematical methods in data science*. Elsevier.

Rodrigues, J., Cancho, V. G., de Castro, M., and Louzada-Neto, F. (2009). On the unification of long-term survival models. *Statistics & Probability Letters*, 79(6):753–759.

Roth, J. A., Yuan, Y., Othus, M., Danese, M., Wagner, S., Penrod, J. R., and Ramsey, S. D. (2021). A comparison of mixture cure fraction models to traditional parametric survival models in estimation of the cost-effectiveness of nivolumab for relapsed small cell lung cancer. *Journal of Medical Economics*, 24(1):79–86.

Schick, A. and Yu, Q. (2000). Consistency of the gmle with mixed case interval-censored data. *Scandinavian Journal of Statistics*, 27(1):45–55.

Schneider, K. A. (2018). Large and finite sample properties of a maximum-likelihood estimator for multiplicity of infection. *PloS one*, 13(4):e0194148.

Scolas, S., El Ghouch, A., Legrand, C., and Oulhaj, A. (2016). Variable selection in a flexible parametric mixture cure model with interval-censored data. *Statistics in Medicine*, 35(7):1210–1225.

Seppä, K., Hakulinen, T., Kim, H.-J., and Läärä, E. (2010). Cure fraction model with random effects for regional variation in cancer survival. *Statistics in medicine*, 29(27):2781–2793.

Shireman, E., Steinley, D., and Brusco, M. J. (2017). Examining the effect of initialization strategies on the performance of gaussian mixture modeling. *Behavior research methods*, 49:282–293.

Singla, N., Jain, K., and Sharma, S. K. (2012). The beta generalized weibull distribution: properties and applications. *Reliability Engineering & System Safety*, 102:5–15.

Statisticat, L. (2013). Bayesian inference. *Farmington, CT: Statisticat, LLC*.

Stephens, M. A. (1974). Edf statistics for goodness of fit and some comparisons. *Journal of the American statistical Association*, 69(347):730–737.

Su, C.-L., Chiou, S. H., Lin, F.-C., and Platt, R. W. (2022). Analysis of survival data with cure fraction and variable selection: A pseudo-observations approach. *Statistical methods in medical research*, 31(11):2037–2053.

Sy, J. P. and Taylor, J. M. (2000). Estimation in a cox proportional hazards cure model. *Biometrics*, 56(1):227–236.

Taweab, F. A. (2015). Parametric cure fraction models for interval–censoring with a change–point based on a covariate threshold.

Tawiah, R., McLachlan, G. J., and Ng, S. K. (2020). A bivariate joint frailty model with mixture framework for survival analysis of recurrent events with dependent censoring and cure fraction. *Biometrics*, 76(3):753–766.

Terwilliger, T. and Abdul-Hay, M. (2017). Acute lymphoblastic leukemia: a comprehensive review and 2017 update. *Blood cancer journal*, 7(6):e577–e577.

Tsodikov, A. (1998). A proportional hazards model taking account of long-term survivors. *Biometrics*, pages 1508–1516.

Tsodikov, A. (2001). Estimation of survival based on proportional hazards when cure is a possibility. *Mathematical and Computer modelling*, 33(12-13):1227–1236.

Tsodikov, A. (2002). Semi-parametric models of long-and short-term survival: an application to the analysis of breast cancer survival in utah by age and stage. *Statistics in medicine*, 21(6):895–920.

Tsodikov, A., Ibrahim, J. G., and Yakovlev, A. (2003). Estimating cure rates from survival data: an alternative to two-component mixture models. *Journal of the American Statistical Association*, 98(464):1063–1078.

Tsodikov, A. D., Yakovlev, A. Y., and Asselain, B. (1996). *Stochastic models of tumor latency and their biostatistical applications*, volume 1. World Scientific.

Tucker, L. and Jmg Taylor, S. (1996). Improved models of tumour cure. *International journal of radiation biology*, 70(5):539–553.

Usman, U., Shamsuddeen, S., Arkilla, B. M., and Yakubu, A. (2022). Mixture cure model for right censored survival data with weibull exponentiated exponential distribution. *Pakistan Journal of Statistics*, 38(4).

Wang, P., Zhao, H., and Sun, J. (2016). Regression analysis of case k interval-censored failure time data in the presence of informative censoring. *Biometrics*, 72(4):1103–1112.

Wei, L.-J., Lin, D. Y., and Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *Journal of the American statistical association*, 84(408):1065–1073.

Weinberg, M. D. (2013). Computational statistics using the bayesian inference engine. *Monthly Notices of the Royal Astronomical Society*, 434(2):1736–1755.

Wilkinson, D. J. (2007). Bayesian methods in bioinformatics and computational systems biology. *Briefings in bioinformatics*, 8(2):109–116.

Xu, J. and Peng, Y. (2014). Nonparametric cure rate estimation with covariates. *Canadian Journal of Statistics*, 42(1):1–17.

Yakovlev, A. Y., Asselain, B., Bardou, V., Fourquet, A., Hoang, T., Rochefediere, A., and Tsodikov, A. (1993). A simple stochastic model of tumor recurrence and its application to data on premenopausal breast cancer. *Biometrie et analyse de donnees spatio-temporelles*, 12:66–82.

Yamaguchi, K. (1992). Accelerated failure-time regression models with a regression model of surviving fraction: an application to the analysis of "permanent employment" in japan. *Journal of the American Statistical Association*, 87(418):284–292.

Yang, J., Li, Y., Liu, Q., Li, L., Feng, A., Wang, T., Zheng, S., Xu, A., and Lyu, J. (2020). Brief introduction of medical database and data mining technology in big data era. *Journal of Evidence-Based Medicine*, 13(1):57–69.

Yin, G. and Ibrahim, J. G. (2005a). Cure rate models: a unified approach. *Canadian Journal of Statistics*, 33(4):559–570.

Yin, G. and Ibrahim, J. G. (2005b). A general class of bayesian survival models with zero and nonzero cure fractions. *Biometrics*, 61(2):403–412.

Yiqi, B., Cancho, V. G., Dey, D. K., Balakrishnan, N., and Suzuki, A. K. (2020). Power series cure rate model for spatially correlated interval-censored data based on generalized extreme value distribution. *Journal of Computational and Applied Mathematics*, 364:112362.

Yiqi, B., Cancho, V. G., Louzada, F., and Suzuki, A. K. (2017). Cure rate proportional odds models with spatial frailties for interval-censored data. *Communications for Statistical Applications and Methods*, 24(6):605–625.

Yu, B. and Peng, Y. (2008). Mixture cure models for multivariate survival data. *Computational Statistics & Data Analysis*, 52(3):1524–1532.

Zeng, D., Yin, G., and Ibrahim, J. G. (2006). Semiparametric transformation models for survival data with a cure fraction. *Journal of the American Statistical Association*, 101(474):670–684.

Zhang, J. and Peng, Y. (2009). Accelerated hazards mixture cure model. *Lifetime data analysis*, 15:455–467.

Zhang, Z. and Sun, J. (2010). Interval censoring. *Statistical methods in medical research*, 19(1):53–70.

Zhao, X. and Curtis, A. (2024). Bayesian inversion, uncertainty analysis and interrogation using boosting variational inference. *Journal of Geophysical Research: Solid Earth*, 129(1):e2023JB027789.

Zyphur, M. J. and Oswald, F. L. (2015). Bayesian estimation and inference: A user's guide. *Journal of Management*, 41(2):390–420.