An overview of time-to-event analysis in dental research

Tabitha K. Peter, Erliang Zeng

November 22, 2022

Abstract

Time-to-event (survival) analysis is an integral tool in the wheelhouse of the dental researcher. While there are many references available for the study of time-to-event analysis, these references tend to be written for audiences trained in statistical methodology. Moreover, the canonical examples offered by most time-to-event analysis references are focused on outcomes which do not translate directly to dentistry. In this article, we provide a tutorial of time-to-event analysis written for the specific context of dental research. Our tutorial assumes no statistical training or computing experience. Using real data from a dental study as our extended example, we explain foundational concepts, including median survival, Nth year survival, the log-rank test, and the Cox model.

Introduction

The objective of this tutorial is to introduce the foundational concepts of time-to-event analysis for dental researchers. After introducing some key terms, we explain the essential concepts of time-to-event analysis using a data set from the dental research literature. See this link for details of this data set that was included in a publication (Jain et al. (2022)).

Word bank

We begin our exploration of time-to-event analysis by highlighting some key terms. Detailed definitions of these terms are provided in the Dictionary section. Notice that several of these words are part of the common vernacular, but have a particular meaning within the context of time-to-event analysis.

Extended example with dental data

In this extended example, we will analyze a data set with time-to-event information on a set of crown margin repairs. Over the course of 10+ years, 1,002 patients received a crown margin repair treatment at the University of Iowa Dental Clinic. Data were collected for each patient via chart review from electronic dental records. The details of data collection and inclusion/exclusion criteria are available in the original manuscript (Jain et al. (2022)).

Our objective for this tutorial is to analyze how long the crown margin repairs lasted. Our observations are the 1,002 crown margin repairs (CMRs), each representing a unique patient. An 'event' is a documented re-intervention on the CMR (such as replacement or extraction). CMRs that are not documented as having events during the study timeframe are the censored observations. The survival time for each CMR is the time (in years) between the date that the patient received the CMR treatment and the last date that the CMR was documented in the data.[1]

[1]The analysis presented in this tutorial differs from what is presented in the publication. This is because the primary objective of the approach presented here is didactic.

We are interested in studying how several factors impact the lifespan of these CMRs. The factors we study here include:

- Age: patient's age (years)[2]
- Sex: patient's sex
- Caries Risk Assessment (CRA): the caries risk of the tooth that received the CMR
- Tooth Type: whether the CMR was placed on an anterior (A) or posterior (P) tooth
- Jaw: whether the CMR was placed in the maxilla (Mx) or mandible (Md)

- Repair material: the material used for the CMR
- Surfaces: the type of surface on which the CMR was done (Buccal, lingual, or other)
- Number of surfaces: how many surfaces were involved in the CMR
- Root canal treatment (RCT): was the tooth receiving the CMR root canal treated?
- Crown type: the type of crown used for the CMR
- Provider: whether the provider placing the CMR was a faculty member or a student

Descriptive analysis

Once we have verified that the data set is ready for analysis, we create a table that summarizes each variable in the data set. In Table 2, numeric values (like Age) are summarized with their median and range (min, max) values. Categorical variables are summarized by their counts and percentages.

Timeline chart

To illustrate the stories of specific patients, we drew a timeline chart (some people call this type of graph a "swimlane diagram"). This kind of diagram lets us put a loupe on our data and see what is going on for an individual patient.

The timeline chart tells the stories of five patients from the CMR study. Patient 438 had a CMR repair that lasted for more than six years. During those six years, the patient had dental visits at the UI Dental Clinic, and so the researchers could verify that the patient's CMR was intact. After those 6+ years, we do not know what happened to this patient - there is no more data available for that person's CMR. Perhaps this person moved out of town, or started going to another clinic. In this case, Patient 438 is counted as 'censored', meaning this person's CMR did not need a re-intervention during the time that the patient was part of the study. Similarly, Patient 359 and Patient 1 are also censored, meaning that neither of their CMRs needed a re-intervention during the time of the study. Patient 17 has a censoring mark (the dark circle) at time 0 - this means that Patient 17 received a CMR treatment from the UI Dental Clinic during the time that the researchers were collecting data, but never returned to that clinic again. The researchers at UI do not have any information about Patient 17 beyond the CMR treatment, and so this patient is 'censored at baseline.'

Patient 150 is the only patient in the timeline chart that has an 'event' marker, indicating that this patient's CMR needed a re-intervention after six years. Maybe this patient needed an extraction, or the teeth involved in the CMR required treatment for caries. This falls within the definition of 'event' that the researchers chose before analyzing the data for this study.

Kaplan-Meier plot

To get an idea of the overall trajectory of the time that the crown margin repairs are lasting, we may examine a Kaplan-Meier (KM) plot (Figure 2). As time goes on, there are fewer repairs upon which to draw estimates, so our estimates become more uncertain at later years. To visualize this uncertainty, we can add confidence intervals to the plots. A risk table provides details to supplement the general pattern illustrated in the KM plot.

[2]The reader may notice in the descriptive table that there are a couple of patients who are > 100 years old. We confirmed with the clinicians that this is correct.

Whereas the timeline chart illustrated the stories of individual patients, the KM plot communicates the overarching story of the entire group of patients. The vertical ('y') axis of the KM plot shows the proportion (fraction) of CMRs that have **not** yet required a re-intervention. This value can be interpreted as the probability that a repair is intact. The horizontal ('x') axis shows time, where year 0 is the time that patients received the CMR treatments. The black line that looks like a staircase is the KM curve, representing the KM estimates of survival probability at each time. The colored area around that black line represents the 95% confidence interval.

At year 0, all patients have a CMR intact, and so the proportion on the vertical axis is at 1. We say that the probability of having an intact CMR at year 0 is 100%. As time moves onward (from left to right), some patients start to need re-interventions. At each re-intervention, the KM curve drops down, giving it the 'staircase' appearance. The downward path of the KM curve illustrates that the probability of a CMR being intact is decreasing over time. By year 5, about half of the CMRs have required re-interventions. By year 10, the probability of a repair being intact is less than 25%.

Simultaneous with the downward path of the KM curve is the widening of the colored area. In the first two years, the colored area keeps tightly around the KM curve; however, by year 10 the colored area is quite widespread. This illustrates that as more CMRs require re-interventions, there is increasing uncertainty about the Kaplan-Meier estimate of survival probability. Generalizing beyond our example, confidence intervals typically get wider over time, reflecting the uncertainty in the KM survival probability estimation.

The risk table beneath the KM plot is lined up with the horizontal axis, indicating that the risk table information is also dependent on time. At each time point, the risk table shows two values: the number of CMRs 'at risk' at that time (top number), and the number of CMRs that have had 'events' up to that time (bottom number). "At risk" means the number of observations which have 1) not yet had an event and 2) have not yet been censored. "Events" indicates the total number of observations that have had events up to a specific time.

At year 0, all 1,002 CMRs are at risk. We do notice that four CMRs are marked as having events in year 0 - at first, this does not make clinical sense. Such a data phenomena is often a sign of a bookkeeping issue, and that is the case here. For these four individuals, the researcher doing the chart review/data extraction should go back and read the clinical notes to determine the best way to document what happened to the patients represented by these four CMRs. In a typical time-to-event analysis, the number of events at the baseline time is 0.

Moving forward in time, we see that at 2.5 years there are 188 CMRs which have already required re-interventions. A total of 421 of our initial 1,002 CMRs are still at risk at this time, meaning we are still collecting data from the patients with these CMRs and none of them have required a re-intervention yet. Notice that 188 + 421 does not add up to 1,002 – the other 393 CMRs are neither at risk nor have they required re-intervention, which means they have been *censored* by year 2.5.

By year 12.5, there are no more observations at risk, meaning that all the CMRs have either been censored or required re-interventions.

Analysis

Median survival

In most time-to-event analyses, the authors report the median survival time. This is defined as the time by which approximately half of all the observations have had an event. For our CMR example, the median survival time is the number of years by which approximately half of the CMRs required a re-intervention.

Keep in mind: The median survival time is **not** just the median of all the survival time values. When we are talking about median survival, we have to account for the fact that some repairs are censored – only events that happen during our study can tell us something definitive about survival time! The Kaplan-Meier method for calculating median survival takes censoring into account. [3]

[3] We chose to do this calculation in 'R' with the 'survfit' function from the 'survival' package.

Table 3 shows that the median survival time for the CMRs was 5.11 years, with a 95% confidence interval of 4.48 - 5.72 years. Looking back at our KM plot, this estimate makes sense - we observed that the KM curve was at about 0.5 on the vertical axis at the year = 5 mark.

Nth year survival

We are also often interested to estimate the survival probability of a repair making it N number of years (N can be any number the researcher's choice). Table 4 that shows estimates of 1 year, 3 year, and 5 year survival for our CMRs.

Plot 3 year survival probability and median survival time

To show the difference between estimating N year survival and estimating the median survival time, we can illustrate how these two measurements align with the KM plot. In Figure 3 (below), we show the median survival time and 3 year survival measurements with colored lines. The arrows on the colored lines show where the researcher would start to estimate each of these measurements. For median survival, one begins with survival probability (on the vertical axis). For 3 year survival, one begins with time (on the horizontal axis).

To keep this illustration from being too visually 'busy', we leave off the confidence intervals from the KM plot.

Comparing groups

Up to this point, we have been studying a Kaplan-Meier plot that describes the entire data set (all crown margin repairs). In practice, the objective is often to compare two subgroups from within the data set – for instance, suppose we are interested in comparing how well crown margin repairs lasted between the root canal treated (RCT) and non-RCT groups. Figure 4 draws two Kaplan-Meier survival curves – one for each of these subgroups. We notice that across time, the curve representing the RCT teeth is consistently below the curve representing the non-RCT teeth. This indicates that the curve for the RCT teeth is *dropping (decreasing) faster*, illustrating that the crown margin repairs done on RCT treated teeth do not last as long as the repairs done on non-RCT teeth.

In addition to the curves in this graph, we also see the confidence intervals at each time point illustrated by the shaded area around each curve. The yellow and purple tinted areas overlap with each other quite a bit, which symbolizes that the difference between crown margin repairs done on RCT teeth and non-RCT teeth is subtle – the repairs last only slightly longer on the non-RCT teeth. Corresponding to these observations, the p-value from the log-rank test shows that this difference is significant at the $\alpha = 0.05$ level.

As a second example of comparing Kaplan-Meier plots between groups, let us suppose that we are working in a materials science context, where we are interested in comparing crown margin repairs that were done with glass ionomer (GI) to repairs done resin-modified glass ionomer (RMGI). We see in Figure 5 that the survival curve representing the GI group is much lower than the curve for the RMGI group for all times after two years. We also see that the space between the two curves increases over time - the two curves are diverging. The confidence intervals do not overlap much at all after 2.5 years. These survival curves indicate that the crown margin repairs done with RMGI lasted notably longer than the repairs done with GI. There is evidence in this data set that the modification to GI makes a positive impact on the expected lifespan of crown margin repairs. Again, corresponding to these observations, the p-value from the log-rank test shows that this difference is notably significant at the $\alpha = 0.05$ level.

Log-rank test In Figures 4 & 5, the p-value shown on the graph is the result of a **log-rank test**. The log-rank test is a popular test in time-to-event (survival) analysis settings where the goal is to compare two or more groups of observations *without* controlling for any other factors. As a statistical tool, the log-rank test examines the data to see if there is evidence against this null hypothesis: "there is no difference in survival between the groups being compared."[4] In other words, the log-rank test can be thought of as a test of whether the survival curves for the groups are identical (overlapping) or not (LaMorte (2016)). For Figure

[4] i.e. it is like the Cohran-Mantel-Hanzel (CMH) test for categorical data.

4, the log-rank test is evaluating whether the survival curves for the RCT and non-RCT CMRs are the same, and there is evidence ($p = 0.009$) against the null hypothesis – that is, there is evidence that the RCT and non-RCT CMRs do not have the same survival curves. Similarly, Figure 5 shows evidence from a log-rank tests that CMRs done with RMGI and GI do not have the same survival curve ($p < 0.001$).

There are several variations of the log-rank test that may be better for comparing groups in certain research settings. Some variants of the log-rank test focus on identifying differences between survival curves earlier in the follow-up time, while other variants focus on differences later in the follow-up time. Consult a biostatistician during the design phase of a proposed study in order to discern if one of the variations of the log-rank test is appropriate.

Cox model Our final survival analysis tool for this tutorial is a Cox proportional hazards model. This model is designed to examine the impact of **multiple factors** in relationship the time-to-event outcome. In practice, the Cox models is often used when there is one particular grouping or factor of interest and several other factors for which the investigator wants to control. Using our example data, we fit a Cox model that examines the relationship between repair material and the time until re-intervention while controlling for age, sex, CRA status, tooth type, jaw, number of surfaces, RCT status, crown type, and provider type.

Table 5 summarizes the results of this Cox model using hazard ratios (HR), 95% confidence intervals (CI), and p-values. Hazard ratios (as defined in the Dictionary) indicate the multiplicative impact of an independent variable on the probability of survival. In the popular Cox regression model, the hazard ratio for an independent variable can be calculated by taking the model coefficient (i.e. the β value) for that independent variable and raising e to the power of that β value: $e^{\beta} = \text{HR}$. In Table 4, this step of *exponentiating* model coefficients has already been done. [5]

Now that we have defined hazard ratio, we can interpret some of the hazard ratios from Table 4:

- For repair material glass ionomer (GI), the hazard ratio is 2.59, with a corresponding p-value of 0.008. This result indicates that at any given year, a CMR done with GI is 2.59 times as likely to need a re-intervention than a CMR done with Amalgam (Amal, the reference category). The impact of using GI as the repair material is notably significant at the $\alpha = 0.05$ level.
- For root canal treatment (RCT), the hazard ratio indicates that CMRs placed on root canal treated teeth are 1.56 times as likely to need a re-intervention compared to CMRs placed on non-RCT teeth. The impact of RCT is mildly significant (or 'suggestive') at the $\alpha = 0.05$ level.
- For jaw, the hazard ratio is 0.67 for maxilla (Mx), indicating that CMRs placed in the maxilla are 0.67 times as likely to require re-intervention compared to CMRs placed in the mandible (the reference category). The impact of placing a CMR in the maxilla is not significant at the $\alpha = 0.05$ level.

From these examples, notice that hazard ratios bigger than 1 correspond to an increased probability of an event, whereas hazard ratios smaller than 1 correspond to a decreased probability of an event.

Summary

In this tutorial, we have studied how to describe and analyze time-to-event ('survival') data using an example from dental research. The tools we used to describe data included timeline charts and Kaplan-Meier plots. To analyze data, we used median survival time, Nth year survival probability, risk tables, and a Cox proportional hazards model. This set of tools provides the dental researcher a place to begin when designing, conducting, or studying time-to-event data.

Dictionary

- **Observation**: The unit of study. These units could be people, as in the case of a clinical investigation where each unit of study is a patient. An observation could also be a dental implant, a set of dentures,

[5] To see what 'R' code I used to format the table this way, refer to the 'analysis.Rmd' files in the GitHub page for this tutorial.

or a plate of bacteria. Regardless of the research context, the observation is studied by the researcher for a specified amount of time. For each observation, there is a time of entry into the study and a time of last observation. In addition to these dates/times, information about the details of an observation are also recorded.

- **Time**: The units of time that have passed since an observation's entry into the study. In time-to-event analysis, the date/time of entry into the study is labeled as 0, and the subsequent units of time (e.g. days, weeks, years) count forward from that starting time. For example, suppose the observations in my study are patients treated in a specific dental clinic, and I am recording time in years. If Patient A enters the study in 2018 and I study that patient until 2022, then I would call 2018 "year 0" and 2022 "year 3." Note that the time of entry into the study often differs among observations. Returning to our example, suppose Patient B enters my study in 2019. In this case, "time 0" for Patient B is 2019. When I make generalizations about all the patients in the study, I would reference "time 0", understanding that this is different calendar years for patients A and B.
- **Event**: The occasion, occurrence, or sign related to the outcome of interest. Every scientific investigation should have an established research objective that informs the choice of an outcome. The choice of this outcome determines the choice of event for a time-to-event analysis. Suppose I want to compare different kinds of dentures to assess which ones are more durable (i.e. which dentures last the longest). In this case, I would need to define what it means for a denture to "last" - perhaps this means that the denture still fits well and does not require replacement. For this context, the outcome of interest could be the amount of time until replacement, and the "event" could be defined as replacement. The observations in the study would be dentures, and the goal would be to study the dentures over time and record the dates and details of those which require replacement. We would say that the dentures which need to be replaced are the 'observations which have an event.'
- **Censored**: The state of observations which do **not** have an event recorded during the time of the study. We describe such observations as "being censored", as opposed to those observations which have events. Censoring can occur when an observation is lost to follow up or does not have an event before the end of a study. Suppose again that I am studying different kinds of dentures to compare their longevity. If a denture is still functioning at the time when I stop collecting data, then I would record this as a censored observation. If a denture is lost to follow up (meaning that the patient with the denture stops coming to my clinic after enrolling in the study), I would also mark that denture as censored. Notice that in both cases, I would not know how long the denture lasted. For all dentures in my study, the last time I collected data about a denture is the date at which I record that denture as 'being censored' or 'having an event'.
- **Survival time**: The time elapsed from the time of entry into the study until the time of either an event or censoring. Consider a denture that I begin studying in 2013 and which needs a replacement in 2021. I would say this denture had an event, and it had a survival time of 8 years.
- **Median survival time**: The time by which approximately half of all the observations in a study had experienced an event. For example, suppose I am studying 50 patients with full dentures, and the event of interest is replacing those dentures. If after 9 years, 25 of those dentures have been replaced, I would report 9 years as the median survival time.
- **Nth year survival (e.g. 3 year survival)**: The proportion of observations that have **not** had an event after 'n' units of time (where 'n' can be any number). Back to the denture example, suppose I am interested in studying 10 year survival in a data set of 50 dentures. If after 10 years, 15 dentures (30% of 50) have still not been replaced and have not been censored, then I would report 10-year survival as 30%.
- **Kaplan-Meier plot**: This name is a reference to the foundational publication by Kaplan and Meier (1958) and the work generated from its ideas. In brief, Kaplan and Meier presented a method for working with data that has censoring. Kaplan-Meier plots are curves that can be used to illustrate the time-to-event phenomena over time. These plots represent both events and censored observations. We will examine some examples of these curves later in the tutorial.

- **Hazard ratios**: A multiplicative factor used to compare the survival probability between groups. Suppose that in our denture example, we compare dentures made with material A to dentures made with material B. If we find that the hazard ratio for dentures made with material A is 2.5, then the data are showing that at any given time, those dentures made with material A are 2.5 times as likely to need replacement compared to dentures made with material B. Keep in mind that a hazard ratio of 1 means lack of association, a hazard ratio greater than 1 suggests an increased risk, and a hazard ratio below 1 suggests a smaller risk. Dawson, Blanchette, and Pihlstrom (2021) provide an extended explanation of this concept.
- **Log-rank test**: A statistical test used to compare the survival curves between two or more groups. This method does not allow for the controlling of other factors, *i.e.* this is a bivariate method.
- **Cox (proportional hazards) model**: This is a reference to the foundational publication by Cox (1972). In brief, Cox proposed a method for a regression-like analysis which is specially crafted for time-to-event data. "Regression-like" means that the method is useful for considering several independent variables simultaneously, as in the case where one wants to study the effect of one variable while controlling for several other factors. The Cox regression model lets us make generalizations about the impact of independent variables on the outcome of interest using hazard ratios.

Acknowledgements and further reading

- Zaboor (2022) has published a detailed tutorial for survival (time-to-event) analysis in `R`.
- Clark et al. (2003) has published a series of tutorials which are more in-depth than what we provide here

Table 1: Word Bank

Terms	Brief definition
Observation	Unit of study
Time	Time since entry into study
Event	Occasion that marks the outcome
Censored	No event was observed
Survival time	Time elapsed until event/censored
Median survival time	Time when about half of observations had events
Nth year survival	Proportion of observations still 'alive' year N
Kaplan-Meier plot	Curves that tell the story
Hazard ratio	Measure for comparing probability of 'event'
Log-rank test	Tool for comparing groups without controlling for other factors
Cox model	Tool for comparing groups while controlling for other factors

Tables

Table 2: Description of data

Characteristic	**N = 1,002**
Age (yrs)	75 (32, 104)
Unknown	445
Sex	
F	294 (53%)
M	262 (47%)
Unknown	446
CRA	
High Risk	159 (41%)
Not High Risk	230 (59%)
Unknown	613
Tooth type	
A	190 (19%)
P	808 (81%)
Unknown	4
Jaw	
Md	474 (47%)
Mx	528 (53%)
Repair material	
Amal	379 (38%)
GI	114 (11%)
RBC	92 (9.2%)
RMGI	416 (42%)
Unknown	1
Surfaces	
B	403 (40%)
L	214 (21%)
Other	385 (38%)
Number of surfaces	
1	724 (72%)
2	278 (28%)

Table 2: Description of data *(continued)*

Characteristic	**N = 1,002**
Root canal treated?	
No RCT	648 (65%)
RCT	348 (35%)
Unknown	6
Crown type	
C	55 (5.5%)
Other	384 (38%)
PFM	560 (56%)
Unknown	3
Provider	
faculty	356 (36%)
student	646 (64%)
Status	
Censored	673 (67%)
Event	329 (33%)
Time	1.92 (0.00, 12.12)

[1] Formats: Median (range), n (%)

Table 3: Median survival time

Median survival time	95% CI (lower)	95% CI (upper)
5.11	4.48	5.72

Table 4: 1, 3, and 5 year survival

Time	Number at risk	Number of events	Probability of survival	Standard Error	95% CI (lower)	95% CI (upper)
1	690	64	0.923	0.009	0.904	0.941
3	342	161	0.664	0.019	0.628	0.702
5	140	60	0.513	0.023	0.471	0.560

Table 5: Cox model results

Characteristic	HR	95% CI	p-value
Age (yrs)	1.01	0.99, 1.03	0.4
Sex			
F			
M	1.18	0.79, 1.74	0.4
CRA			
High Risk			
Not High Risk	0.71	0.47, 1.06	0.092
Tooth type			
A			
P	1.74	0.92, 3.30	0.091
Jaw			
Md			
Mx	0.67	0.43, 1.06	0.086
Repair material			
Amal			
GI	2.41	1.21, 4.82	0.013
RBC	1.53	0.63, 3.72	0.4
RMGI	1.26	0.79, 2.00	0.3
Surfaces			
B			
L	1.28	0.73, 2.26	0.4
Other	1.42	0.84, 2.40	0.2
Number of surfaces			
1			
2	1.36	0.87, 2.13	0.2
Root canal treated?			
No RCT			
RCT	1.53	1.01, 2.32	0.047
Crown type			
C			
Other	1.29	0.47, 3.58	0.6
PFM	1.20	0.45, 3.18	0.7
Provider			
faculty			
student	1.42	0.85, 2.38	0.2

[1] HR = Hazard Ratio, CI = Confidence Interval

Figures

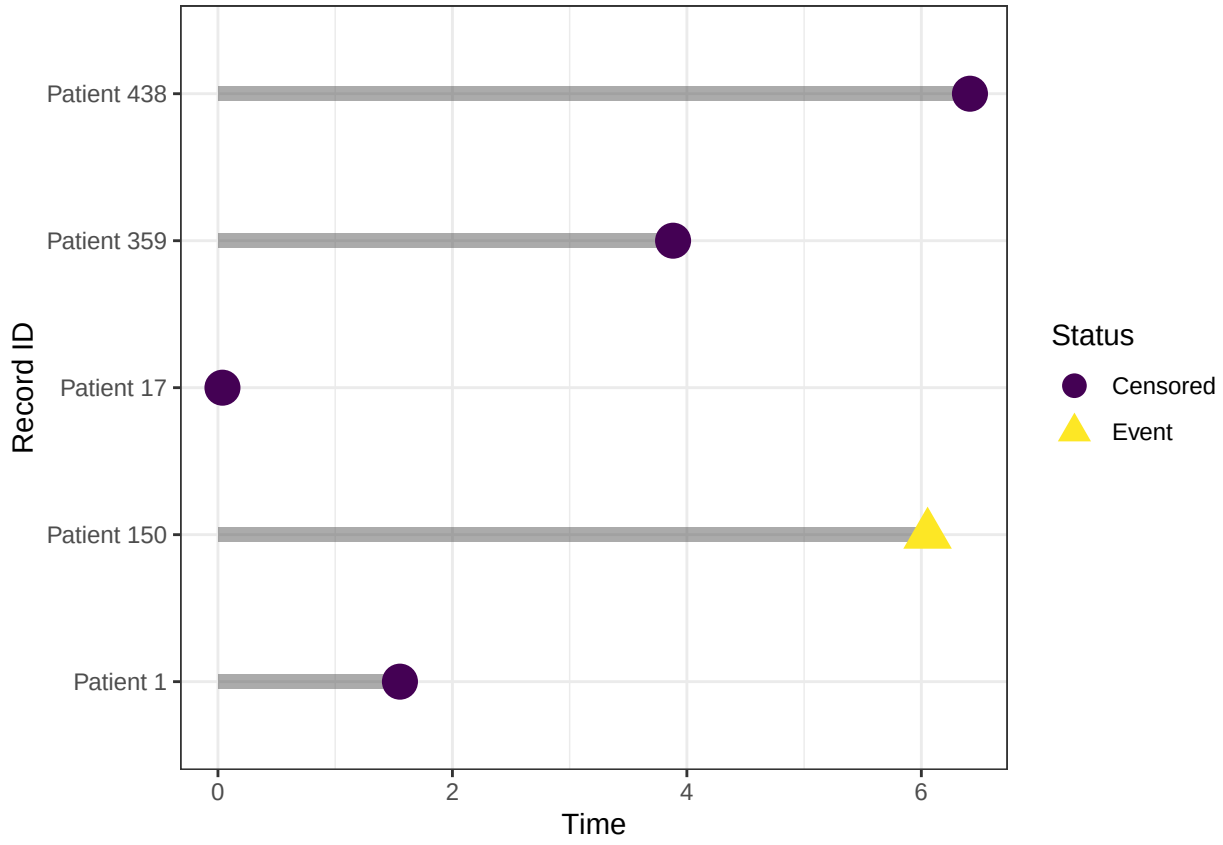


Figure 1: Timeline chart

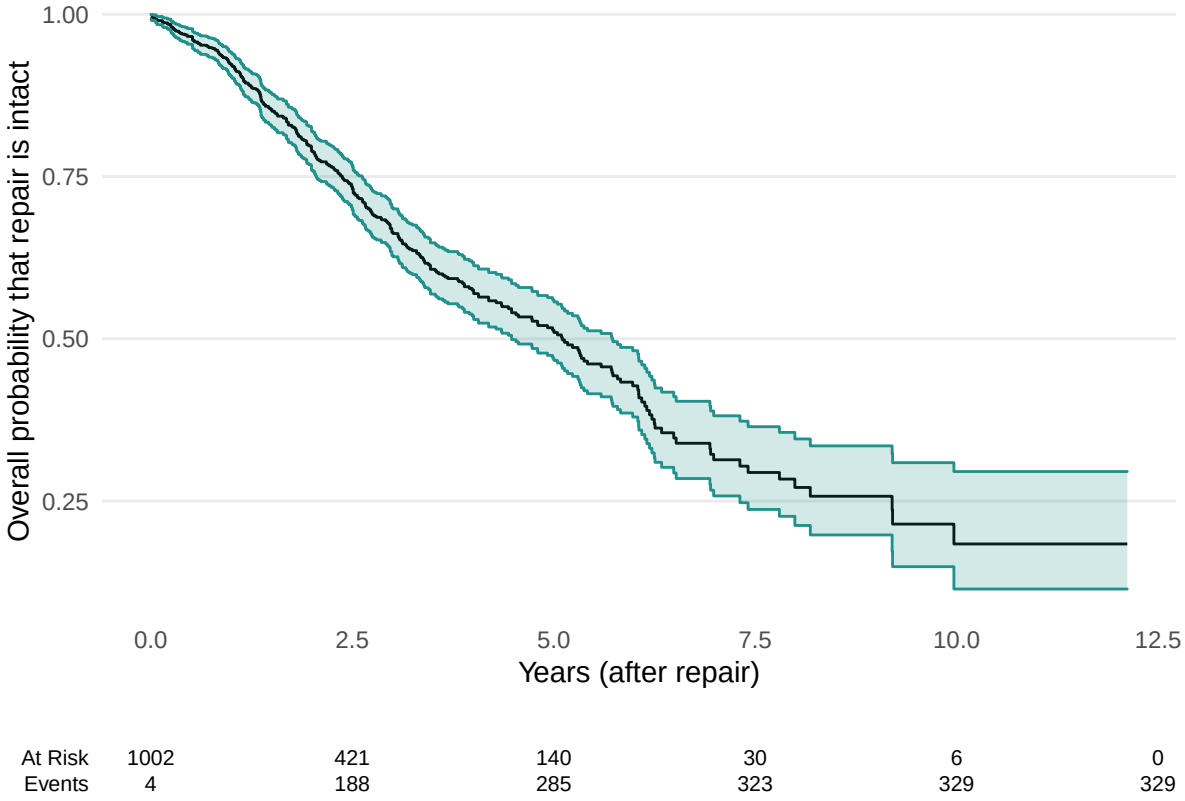


Figure 2: Kaplan-Meier plot

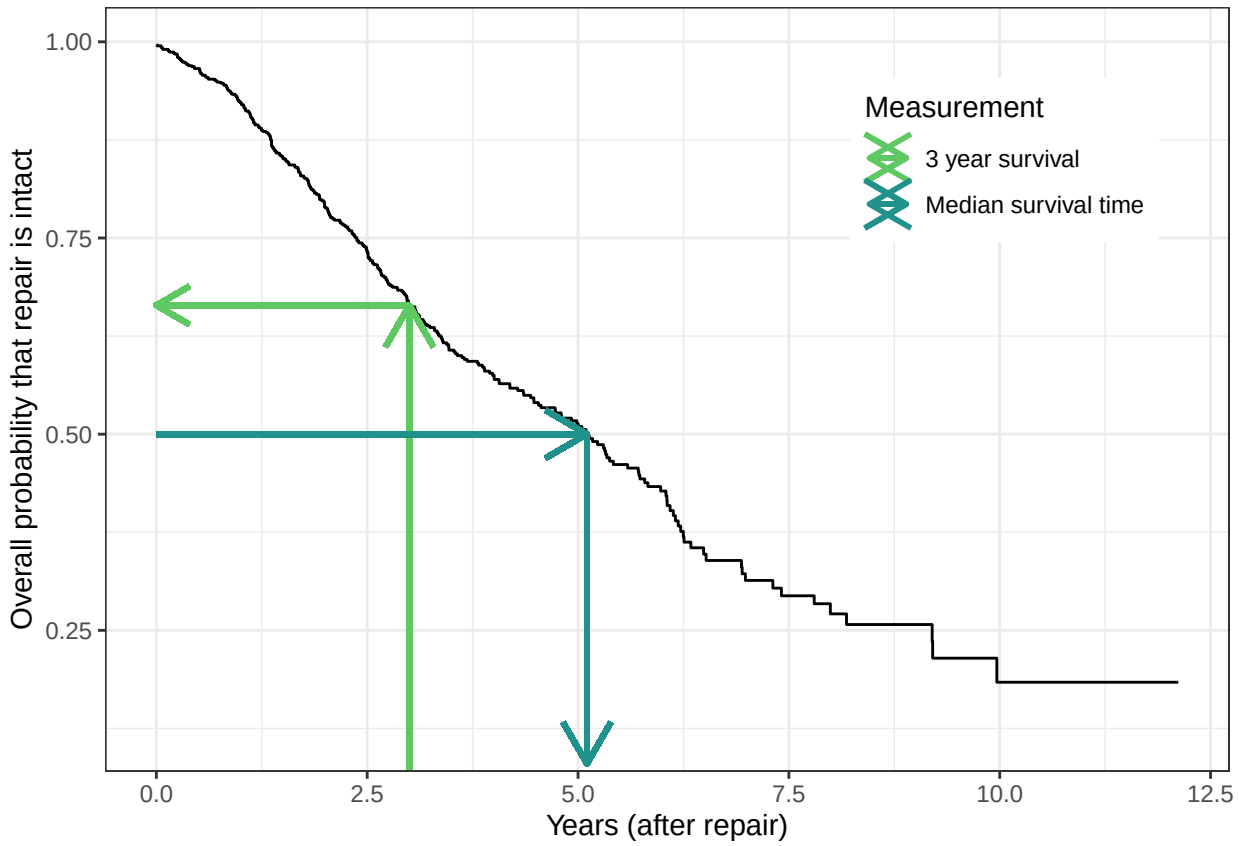
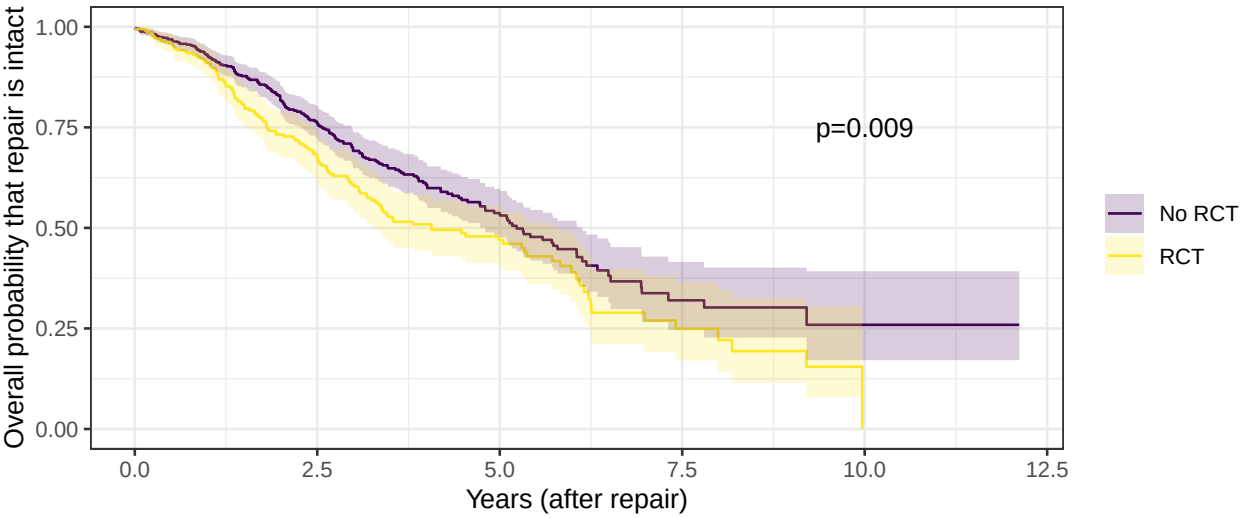
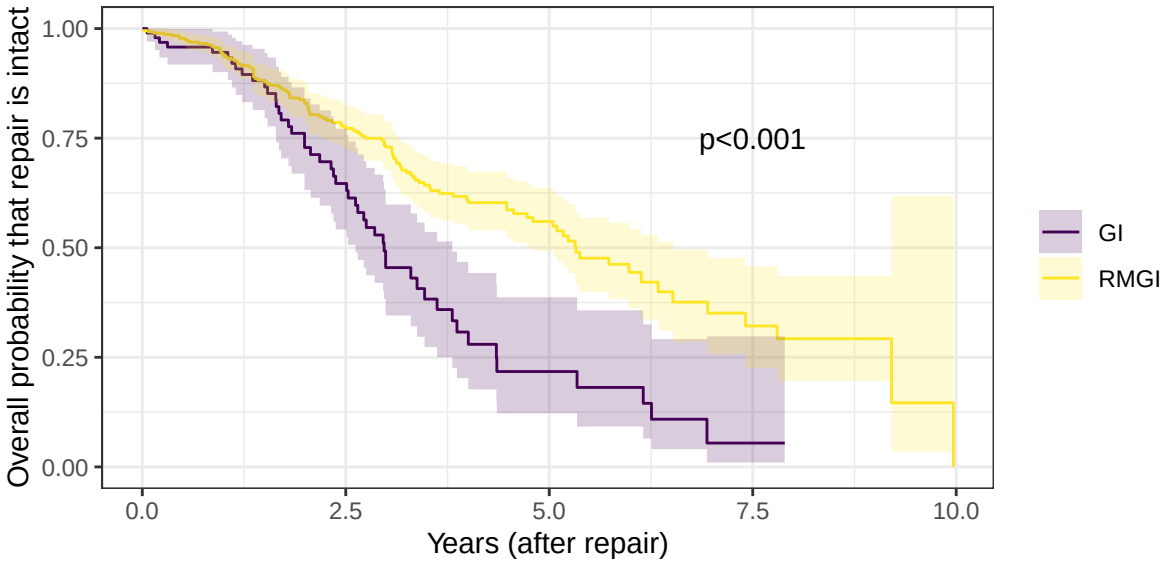


Figure 3: Median survival v. 3-year survival



No RCT						
At Risk	648	283	89	18	6	0
Events	2	110	171	194	196	196
RCT						
At Risk	348	137	51	12	0	0
Events	2	78	113	128	132	132

Figure 4: Comparing RCT groups



GI					
At Risk	114	39	6	1	0
Events	0	25	45	49	49
RMGI					
At Risk	416	174	56	11	0
Events	2	62	98	112	115

Figure 5: Comparing repair materials

References

Clark, Taane G, Michael J Bradburn, Sharon B Love, and Douglas G Altman. 2003. "Survival Analysis Part i: Basic Concepts and First Analyses." *British Journal of Cancer* 89 (2): 232–38.

Cox, David R. 1972. "Regression Models and Life-Tables." *Journal of the Royal Statistical Society: Series B (Methodological)* 34 (2): 187–202.

Dawson, Deborah V, Derek R Blanchette, and Bruce L Pihlstrom. 2021. "Application of Biostatistics in Dental Public Health." In *Burt and Eklund's Dentistry, Dental Practice, and the Community*, 131–53. Elsevier.

Jain, Aditi, Allison Schollmeyer, Tabitha Peter, Xian Jin Xie, and Sindhura Anamali. 2022. "Survival Analysis of Crown Margin Repair: A Retrospective Study in a Dental School Setting." *The Journal of the American Dental Association* 153 (5): 414–20.

Kaplan, Edward L, and Paul Meier. 1958. "Nonparametric Estimation from Incomplete Observations." *Journal of the American Statistical Association* 53 (282): 457–81.

LaMorte, Wayne W. 2016. "Comparing Survival Curves." https://sphweb.bumc.bu.edu/otlt/mph-modules/bs/bs704_survival/BS704_Survival5.html#:~:text=among%20independent%20groups.-,The%20Log%20Rank%20Test,identical%20(overlapping)%20or%20not.

Zaboor, Emily. 2022. "Survival Analysis Tutorial in r." https://www.emilyzabor.com/tutorials/survival_analysis_in_r_tutorial.html.