A Conceptual Approach to Survival Analysis

Laura Lee Johnson, Ph.D.

Associate Director

Division of Biostatistics III

Center for Drug Evaluation and Research

US Food and Drug Administration

IPPCR Course Fall 2015

Chapter 23, 3rd Edition

Disclaimer

This presentation reflects the views of the author and should not be construed to represent FDA's views or policies.

Objectives

“Conceptual Approach”

Recognize some vocabulary used in survival analysis and a few commonly used statistical methods for time to event data in medical research

Interpret Kaplan Meier graph

Interpret a covariate from a Cox model

Understand a few of the assumptions for basic types of survival models

Preventing Mother-Infant HIV Transmission (D.O. Dixon, NIAID)

Zidovudine able to slow progression of HIV in adults with advanced disease

AIDS Clinical Trials Group Protocol 076 designed to assess both safety and efficacy of Zidovudine in preventing transmission of HIV from infected (not advanced) women to their babies

Preventing Mother-Infant HIV Transmission

Powered (80%) to detect a 33% reduction of transmission rate (through 78 weeks of baby's life) relative to projected rate of 30%

Target N was 748; began April 1991

Projected accrual to take at least 5 years and 15% dropouts

Preventing Mother-Infant HIV Transmission

DSMB met twice a year to monitor safety

Efficacy reviews planned after each 1/3 of projected infant infections

1st efficacy review took place in February 1994, based on mothers enrolled up to December 1993 and their babies

P = 0.00006

(Note: Survival Analysis)

What is that graph in the last slide?

Kaplan Meier curves

What are they telling me?

Define the Outcome Variable

What is the event?

What is the y-axis?

Mother-Infant HIV transmission; testing positive

Where is the time origin?

Infant birth?

What is the time scale?

Weeks

How is time at which the event 'occurs' defined?

Kaplan Meier

One way to estimate survival

Or other events (transmission of HIV)

Nice, simple, can compute by hand

Can add stratification factors

No sensible interpretation for competing risks

Cannot evaluate covariates like Cox model

Kaplan Meier is a workhorse

P = 0.00006

Preventing Mother-Infant HIV Transmission

DSMB recommended stopping (after careful review of data quality and completeness, toxicity, transmission rates)

Trial leaders and NIH agreed

Zidovudine provided to those in control group

United States Public Health Service (USPHS) national guidelines modified

Survival or Time-to-Event Analysis

Why Survival (Time to Event) Analysis

New cancer treatment

Want to know if it extends a person's life 5 months longer than current treatments

Survival is about events and when they happen

Death, infection, MI, hospitalization

Recurrence of cancer after treatment

Marriage, soccer goal

Light bulb fails, computer crashes

Balloon filling with air bursts

Why Survival Analysis? Hypertension

Treat it

Lower/Normal blood pressure (BP)

Want to extend healthy life, prevent heart attacks…events in time

Follow people, see how many die and have events like MI and when

What if Intervention lowers blood pressure but after 20+ years people die two years before a similar person not on medicine?

Goal accomplished?

COX 2 inhibitors

People with lower X live longer!

Many times from an observational study

Instigate a change in blood pressure, weight, something

Do you get a similar change in outcome?

Perhaps

People live too long to follow!

We can only hope (and that they live well and feel great, are productive, and all that)

Many surrogates

Even more surrogates for events (in particular death)

Sometimes follow people a long time

What is Survival

Survival analysis deals with making inference about EVENT RATES

Rate at t = Rate among those at risk at t

Look at Median survival (50%) not Mean survival

Mean: need everyone to have an event

Outline

How to Measure Time and Events

Survival and Hazard Functions

Truncation and Censoring

Competing Risks

Models and Hypothesis Testing

Example

Conclusions

What is a Model?

Basic

$Y = \beta 0 + \beta 1 X1 + \ldots + \beta p Xp$

Y = outcome or response variable

β = coefficient

X = covariate, variable

Survival

$\lambda(t) = \lambda 0(t) \exp\{ \beta 1 X1 + \ldots + \beta p Xp\}$

λ0(t) = baseline hazard

β1,…, βp = regression coefficients

X1,…, Xp = prognostic factors

Vocabulary

Survival vs. time-to-event

Outcome variable = event time

Examples of events:

HIV positive test, AIDS defining event, Mother-Infant HIV transmission

Systolic blood pressure or Cholesterol below a cut

BMJ 17 October 2009 article on follow-up of MIST trial: incidence of pregnancy

Time in years to first live birth after index miscarriage

Time Notation

t: for time axis

t = 0 is the time origin

T: random outcome variable

time at which event occurs

Vocabulary

t = time

Baseline = 0 months

6, 12, 18, 24 months, etc.

S(t) = Survival at time t

P[T ≥ t] = Probability Time of event is greater than time t

Define the Outcome Variable

What is the event? (Death?)

Where is the time origin? (Diagnosis?)

What is the time scale? (Weeks? Years?)

Then what?

Need to know the time the event occurs

May or may not use covariates

Could do a logistic regression model

Yes/No outcome

Not focus of lecture

Choice of Time Scale

Scale	Origin	Comment
Study time	Dx or Rx	Clinical Trials
Study time	First Exposure	(Occupational) Epidemiology
Age	Birth (subject)	Epidemiology

Treatment for a Cancer

Event = death

Time origin = date of surgery

Time scale = time (months)

T = time from surgical treatment to death

Graph = $P[T \geq t]$ vs t

Example Numbers

$S(9) = P[T \geq 9] = 0.25$

25% is the probability the time from surgical treatment to death is greater than 9 months

“9 month post-resection survival is 25%”

$0 \leq S(t) \leq 1$

Survival Function

$S(t) = P[T \geq t] = 1 – P[T < t]$

Plot: Y axis = % alive, X axis = time

Proportion of population still without the event by time t

Survival Function in English

Event = death, scale = months since Rx

“S(t) = 0.3 at t = 60”

“The 5 year survival probability is 30%”

“70% of patients die within the first 5 years”

Everyone dies → S(咲 = 0

Hazard Function

Incidence rate, instantaneous risk, force of mortality

λ(t) or h(t)

Event rate at t among those at risk for an event

Key function

Estimated in a straightforward way

Censored

Truncated

Hazard Function in English

Event = death, scale = months since Rx

"λ(t) = 1% at t = 12 months"

"At 1 year, patients are dying at a rate of 1% per month"

"At 1 year the chance of dying in the following month is 1%"

Hazard Function: Instantaneous

120,000 die in 1 year

10,000 die in 1 month

2,500 die in a week

357 die in a day

Instantaneous: move one increment in time

Herpes Example

Recurrence of Herpes Lesions After Treatment for a Primary Episode

Event = recurrence

needs well defined criteria

Time origin = end of primary episode

Time scale = months from end of primary episode

T = time from end of primary episode to first recurrence

Toxin Effect on Lung Cancer Risk

Occupational exposure at nickel refinery

Event = death from lung cancer

Origin = first exposure

Employment at refinery

Scale = years since first exposure

T = time: first employed to death from LC

Population Mortality

Event = death

Time origin = date of birth

Time scale = age (years)

T = age at death

Volume of Air a Balloon Can Tolerate

Event = balloon bursts

t = ml of air infused

Origin = 0 ml of air in the balloon

T = ml of air in balloon when it bursts

Unique Features of Survival Analysis

Event involved

Progression on a dimension (usually time) until the event happens

Length of progression may vary among subjects

Event might not happen for some subjects

Sample Size Considerations

Event may not ever happen for some subjects

Sample sizes based on number of events

Work backwards to figure out # of subjects

Covariates must be considered (age, total exposure, etc)

Outline

How to Measure Time and Events

Survival and Hazard Functions

Truncation and Censoring

Competing Risks

Models and Hypothesis Testing

Example

Conclusions

Truncation and Censoring

Truncation is about entering the study

Right: Only sample those with Event of interest (cancer registry) (underestimate)

Left: short survival may be overlooked (>65 years of age) (overestimate)

Censoring is about leaving the study

Right: Incomplete follow-up (common)

Left: Observed time > survival time (know the subject exists)

Independence is key

Left Truncation

Mention more in epi vs medical studies

Medical: zero-out at time of dx/tx

Key Assumption

Those who enter the study at time t are a random sample of those in the population still at risk at t

Allows one to estimate the hazard function $\lambda(t)$ in a valid way

Who is the audience and will they generalize

Censoring

Incomplete observations

Right

Incomplete follow-up

Common and Easy to deal with

Left

Event has occurred before T0, but exact time is unknown

Not easy to deal with

Left Censoring

Age smoking starts

Data from interviews of 12 year olds

12 year old reports regular smoking

Does not remember when he started smoking regularly

Study of incidence of CMV infection in children

Two subjects already infected at enrollment

One Form of Right Censoring:
Withdrawals

Must be unrelated to the subsequent risk of event for 'independent censoring' to hold

Accidental death is usually ok

Moves out of area (moribund unlikely to move)

Right Censoring

Types of Censoring

Type I censoring

T* same for all subjects

Everyone followed for 1 year

Type II censoring

Stop observation when a set number of events have occurred

Replace all light bulbs when 4 have failed

Random censorship

Our focus, more general than Type I

Notation

T = event time

T* = observation time

T if event occurs

Follow-up time otherwise

∂ = failure indicator

1 if T* = T

0 if T* < T

“censor” or “censor indicator”

Key Assumption:
Independent Censoring

Those still at risk at time t in the study are a random sample of the population at risk at time t, for all t

This assumption means that the hazard function (λ(t)) can be estimated in a fair/unbiased/valid way

Independent Censoring:
If you have Covariates

Censoring must be independent within group

Censoring must be 'independent' given X

Censoring can depend on X

Among those with the same values of X, censored subjects must be at similar risk of subsequent events as subjects with continued follow-up

Censoring can be different across groups

Censoring Examples

At five year follow-up patients have not died

Follow school group from age 5 until 25

Ask students when they start smoking

Answering "never" is an example of right censoring

Early in trial older subjects are not enrolled

Amount of time could be on study?

Condition on age: ok

Do not condition on age: the estimates will be biased because censoring is not independent

Take Away: Study Types

Clinical studies

Time origin = enrollment, treatment begins

Time axis = time on study

Right censoring common

Epidemiological studies

Time axis = age

Right censoring common

Left truncation common

Bottom Line

Standard methods to deal with right censoring and left truncation

Key assumption is that those at risk at t are a random sample from the population of interest at risk at t

Survival Analysis

Models mostly for the hazard function

Accommodates incomplete observation of T

Censoring

Observation of T is ‘right censored’ if we observed only that T > last follow-up time for a subject

Typical Intervention Trial

Accrual into the study over 2 years

Data analysis at year 3

Reasons for exiting a study

Died

Alive at study end

Withdrawal for non-study related reasons (LTFU)

Other reasons, including dying from other causes (which leads us to…)

Outline

How to Measure Time and Events

Survival and Hazard Functions

Truncation and Censoring

Competing Risks

Models and Hypothesis Testing

Example

Conclusions

Competing Risks

Multiple causes of death/failure

Special considerations of competing risk events described in the literature

Example:

event = cancer diagnosis/recurrence

death from myocardial infarction (MI) = competing risk

No basis for believing the independence assumption

Competing Risks

Interpretation of $\lambda(t)$ = "risk of cancer at t when the risk of death from MI does not exist" isn't practically meaningful

Rather, interpret $\lambda(t)$ = "risk of cancer among those at risk of cancer at t"

This will exclude MI deaths (if you are dead from an MI you are not at risk of cancer) and that is ok

Polar Bear Club Death Rates (fiction)

Annual death rates

3% taking dip 1Jan in Lake Michigan

2% Males all other causes

1% Female all other causes

Over a decade

25% of women died from taking a dip in Lake Michigan 1 Jan

24% of men died from taking a dip in Lake Michigan 1 Jan

Polar Bear Club Death Rates (fiction)

Why does it harm women?

Over a decade

33.5% of women died from all other causes

40% of men died from all other causes

There are more women to harm

People die of something

Which means they cannot die from something else

Bottom Line

We make inference about $\lambda_{obs}(t)$ = event rate among subjects under observation at t

We can interpret it as $\lambda(t)$ = event rate among subjects with $T \geq t$, if censoring is independent

Outline

How to Measure Time and Events

Survival and Hazard Functions

Truncation and Censoring

Competing Risks

Models and Hypothesis Testing

Example

Conclusions

Kaplan Meier

One way to estimate survival

Or other events (transmission of HIV)

Nice, simple, can compute by hand

Can add stratification factors

No sensible interpretation for competing risks

Cannot evaluate covariates like Cox model

Kaplan Meier is a workhorse

Kaplan Meier

Multiply together a series of conditional probabilities

Kaplan Meier Curve

Kaplan Meier Estimator

One estimate of S(t)

Use the same idea when looking at 1- S(t)

Need independent censoring

If high risk subjects enter the study late then early on the K-M curve will come down faster than it should

Censored observations provide information about risk of death while on study

What are those little hash marks on the curves? Mark times people were censored

Kaplan Meier

Just the outcome is in many models

One or more stratification variables may be added

Intervention

Gender

Age categories

Quick and Dirty

How to Test? At a Given Time

H0: S1(t) = S2(t)

Form test statistic

“Arbitrary time” – choosing t post hoc

Not using all of the data

Simple Inference

For single event data inference about rates → inference for S(t)

No time dependent covariates, no recurrent events, no competing risk events

Logrank statistics compare event rates and allow the same generality as right censoring, left truncation

Log Rank

H0: S1(.) = S2(.)

Test overall survival

2 independent samples from the same population

Observed # events vs. Expected #

Software; statistician should check

Some variations and some assumptions

Log Rank

Confounding

Are prognostic factors balanced between treatment groups?

Can see a difference using logrank, but just bias

Stratified Log Rank

Compare survival within each stratum

Essentially perform test within each stratum

Can prognostic factor be categorized?

Enough people per stratum?

Loss of power

Significance test, no estimates of difference

P = 0.00006

Cox Models

Proportional Hazards: Cox

Cox Proportional Hazards model

$\lambda(t) = \lambda 0(t) \exp\{ \beta 1 X1 + \ldots + \beta p Xp\}$

$\lambda 0(t)$ = baseline hazard

$\beta 1, \ldots, \beta p$ = regression coefficients

X1,…, Xp = prognostic factors

$\beta = 0 \rightarrow$ hazard ratio = 1

Two groups have the same survival experience

$e\beta$ = relative rate (relative risk) (RR)

Cox Proportional Hazards Model

Add covariates to the model

No need to stratify

Change in a prognostic factor → proportional change in the hazard (on the log scale)

Statistical software

Can test the effect of the prognostic factor as in linear regression - H0: $\beta=0$

Cox Model for Event Rates

Provides a framework for making inference about covariate effects

Semi-parametric

$\lambda_0(t)$ completely unspecified

Multiplicative - $e^{\beta x}$

Effect of covariate is to multiply the rate by a factor

Cox cont.

Requires either that

RR is constant over time (proportional hazards), or

That we model RR over time

Allows time-dependent covariates and stratification factors

Age Example

Early in trial older subjects are not enrolled

If age is not in the Kaplan Meier then the KM estimate is biased because censoring is not independent

Put age in the Cox model – conditioned on age; ok

Age Example (cont.)

If I follow everyone for 1 year, am I ok?

Not necessarily

The study is not proportional by age to the population risk set

Could try to over sample older people later in the study to make the final study more correctly proportional

Easier to condition on age?

Testing Proportional Hazards

$\lambda(t) = \lambda_0(t) \exp\{ \beta_1 \text{ age} + \beta_2 \text{ drug}\}$

$\exp\{ \beta_1 \text{age}+\beta_2 \text{drug}+\beta_3 \text{age}*\ln(t)+\beta_4 \text{ drug } *\ln(t)\}$

Look at p-values associated with β_3 and β_4 (Wald tests)

Do a partial likelihood ratio test comparing the two models

Look at Schoenfeld residual plots

Testing Proportional Hazards

Testing Proportional Hazards

Time-Dependent Survival Curves

Failure to account for change in exposure/treatment over time

Usually assume there is no change

Think about HAART example

Stanford Heart Transplant Study (1971)

End-stage heart disease

Not responding

Seeking transplant

Take Home

Choose the right method and test

Kaplan Meier – simple

Logrank tests – useful, potentially misleading

Cox Proportional Hazards – workhorse

Not everything is proportional – check

Consider more general Cox models

Time matters

Changes in protocol matter

Outline

How to Measure Time and Events

Survival and Hazard Functions

Truncation and Censoring

Competing Risks

Models and Hypothesis Testing

Example

Conclusions

Example

Series of prospective cohort and randomized clinical studies evaluating survival of patients with liver cirrhosis

Compare a ‘new’ treatment, D-penicillamine with placebo

Conflicting reports

One study appears to report a two year survival probability of 0.88, calculated with Kaplan Meier

Trial Information

Data collected at randomization

Lots of information (stage)

Presence/absence of ascites

Prothrombin time in seconds -10

Cox model you might see

$\lambda(t) = \lambda 0(t) \exp\{ -0.135 \text{ XTRT}+1.737 \text{ XA}+0.346 \text{ XP}\}$

What is this model?

$\lambda(t) = \lambda 0(t) \exp\{ -0.135 \text{ XTRT}+1.737 \text{ XA}+0.346 \text{ XP}\}$

XTRT: 1 = D-penicillamine, 0 = placebo

XA: 1 = ascites, 0 = no ascites

XP: Prothrombin time – 10

Continuous, in seconds

$\lambda 0(t)$ is the event rate at time t in the placebo arm for subjects without ascites with a prothrombin time of 10 seconds

$\lambda(t) = \lambda 0(t) \exp\{ -0.135 \text{ XTRT} + 1.737 \text{ XA} + 0.346 \text{ XP} \}$

Relative rate of death two years post randomization for a subject on this trial who received the new treatment, had ascites at randomization and a prothrombin time of 10 seconds compared to a similar subject who received placebo?

$RR = \exp \{ -0.135 \} = 0.87$

Worked Out

RR at Three Years?

Relative rate does not vary with time according to the proportional hazards model.

At the years the previously described RR is also $\exp \{ -0.135 \}$

Can work out RR for lots of other subject comparisons

But…

Physicians were initially reluctant to enter patients with ascites on the trial because of potential toxicity concerns

After about a year and a half recruitment became more representative of the clinic population

Effects on the Validity of the Kaplan Meier (KM) Estimator

Censoring is not independent

At large t, the risk sets will not include patients with ascites because they were not recruited early enough and therefore are censored early.

The hazard function will be biased too small for larger t and so will be larger than the population survival function at large t.

In Short, What If

From first participant entered until the end of study: 4 years

Enroll for 3 years

Can be on study at least 1 year and up to 4 years

Followed enrollment to end of study

Do not start fully enrolling ascites until year 1.5

Ascites Participants

On study at least 1 yr and up to 2.5 yr

Do not have full population/risk set information at time t > 2.5 years

At time points t > 2.5 the study does not include a representative population

Ascites → worse prognosis

KM estimate at t > 2.5 too high

Hazard is too small at larger t

Cox Model: Doomed Regression Coefficient Estimates?

No bias because conditional on covariates (including XA)

Censoring must be independent GIVEN X

Censoring is independent and that is all that is required for consistency of the partial likelihood estimator (i.e. the coefficients)

Outline

How to Measure Time and Events

Survival and Hazard Functions

Truncation and Censoring

Competing Risks

Models and Hypothesis Testing

Example

Conclusions

Survival Analysis

Survival analysis deals with making inference about EVENT RATES

Rate at t = Rate among those at risk at t

Look at Median survival (50%) not Mean survival

Mean: need everyone to have an event

Cox Regression is the most robust method

Kaplan Meier curves do not have sensible interpretations for competing risks

Survival Analysis Can Handle

Right censoring

Left truncation

Recurrent events

Competing risks, etc.

Available representative risk sets at t allow us to estimate/model event rates

Kaplan Meier

One way to estimate survival

Nice, simple, can compute by hand

Can add stratification factors

Cannot evaluate covariates like Cox model

No sensible interpretation for competing risks

Inference: Log Rank

Logrank statistics compare event rates and allow the same generality as right censoring, left truncation

For single event data inference about rates → inference for S(t)

No time dependent covariates, no recurrent events, no competing risk events

Cox Model for Event Rates

Provides a framework for making inference about covariate effects

Semi-parametric

$\lambda 0(t)$ completely unspecified

Multiplicative - $e\beta x$

Effect of covariate is to multiply the rate by a factor

Cox cont.

Requires either that

RR is constant over time (proportional hazards), or

That we model RR over time

Allows time-dependent covariates and stratification factors

Truncation and Censoring

Independence is key

Truncation is about entering the study

Right: Event has occurred (e.g. cancer registry)

Left: Have the event and fall out of view before they can enter to be counted

Censoring is about leaving the study

Right: Incomplete follow-up (common)

Left: Observed time > survival time

Analysis Follows Design

Questions → Hypotheses →

Experimental Design → Samples →

Data → Analyses →Conclusions

Take all of your design information to a statistician early and often

Guidance

Assumptions

Questions?

Thanks!

Please fill out the course evaluations

Post specific examples

Send questions to the appropriate board for each lecture