EMORY
ROLLINS SCHOOL OF PUBLIC HEALTH

Department of Biostatistics and Bioinformatics

BIOS 522: Survival Analysis Methods

Lecture 12:

Competing risks

Competing risks

- **More than one event type** of interest
- Or, there is another type of event that can prevent you from observing the event of interest

- We cannot treat the other events as **censoring times** because the events may be **correlated/dependent/informative**
- Treating these other events as censoring times induces **bias**!

- A simple solution is to define a **composite endpoint** (e.g. death due to any cause) but then we cannot isolate a particular event

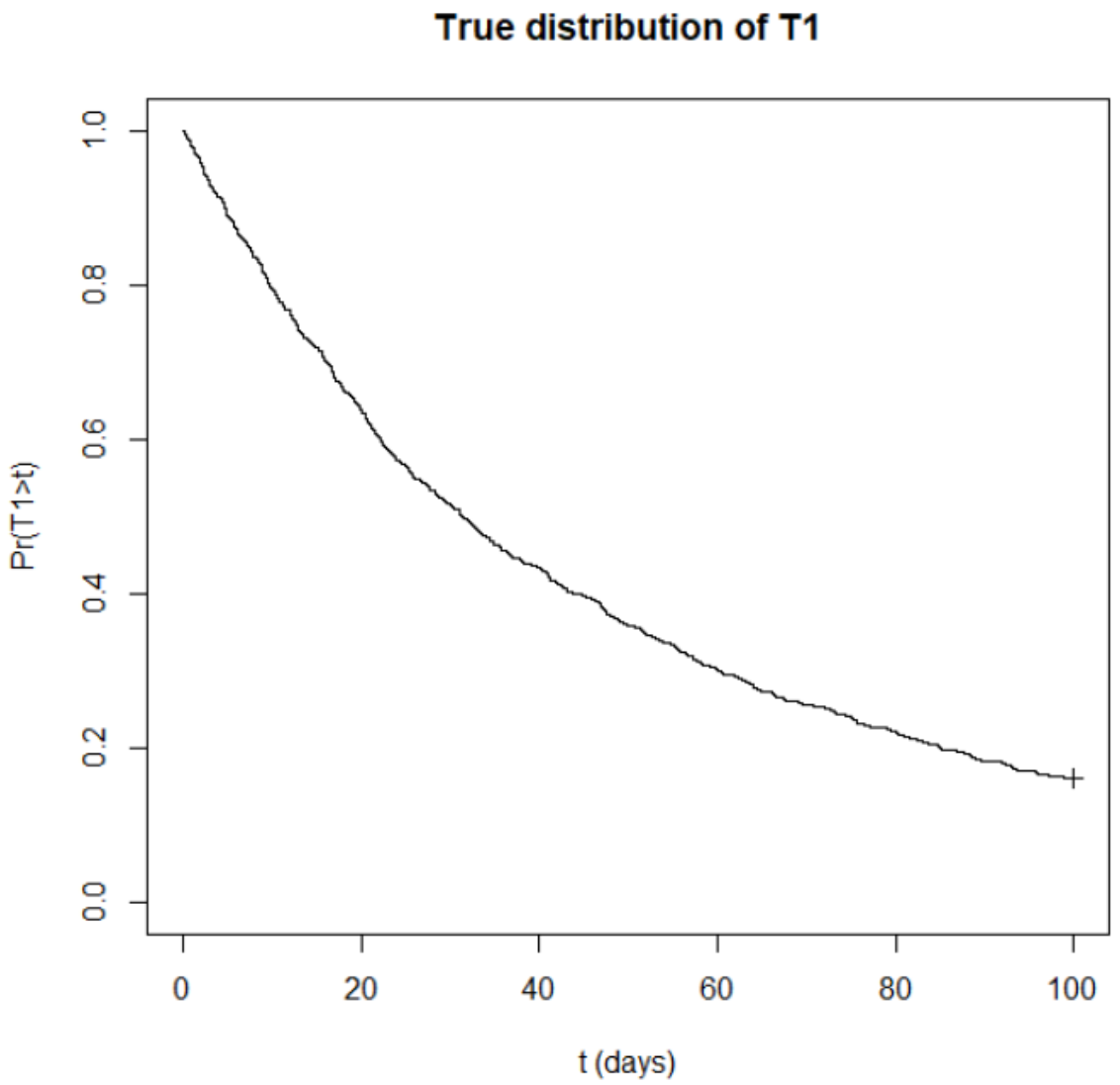
Analysis framework

- Suppose we have m distinct failure types, for example:
  - $m = 1$: cardiovascular death
  - $m = 2$: non-cardiovascular death
- Individuals have hypothetical times to failure:
  - T_1 is the time to cardiovascular death
  - T_2 is the time to non-cardiovascular death
- We observe only the earliest failure time, as well as its type
  - $T = \min(T_1, T_2), \delta = (1,2)$
  - This time can also be censored $T^*, \delta = 0$

Target estimand

- Interested in studying T_1, time to cardiovascular death

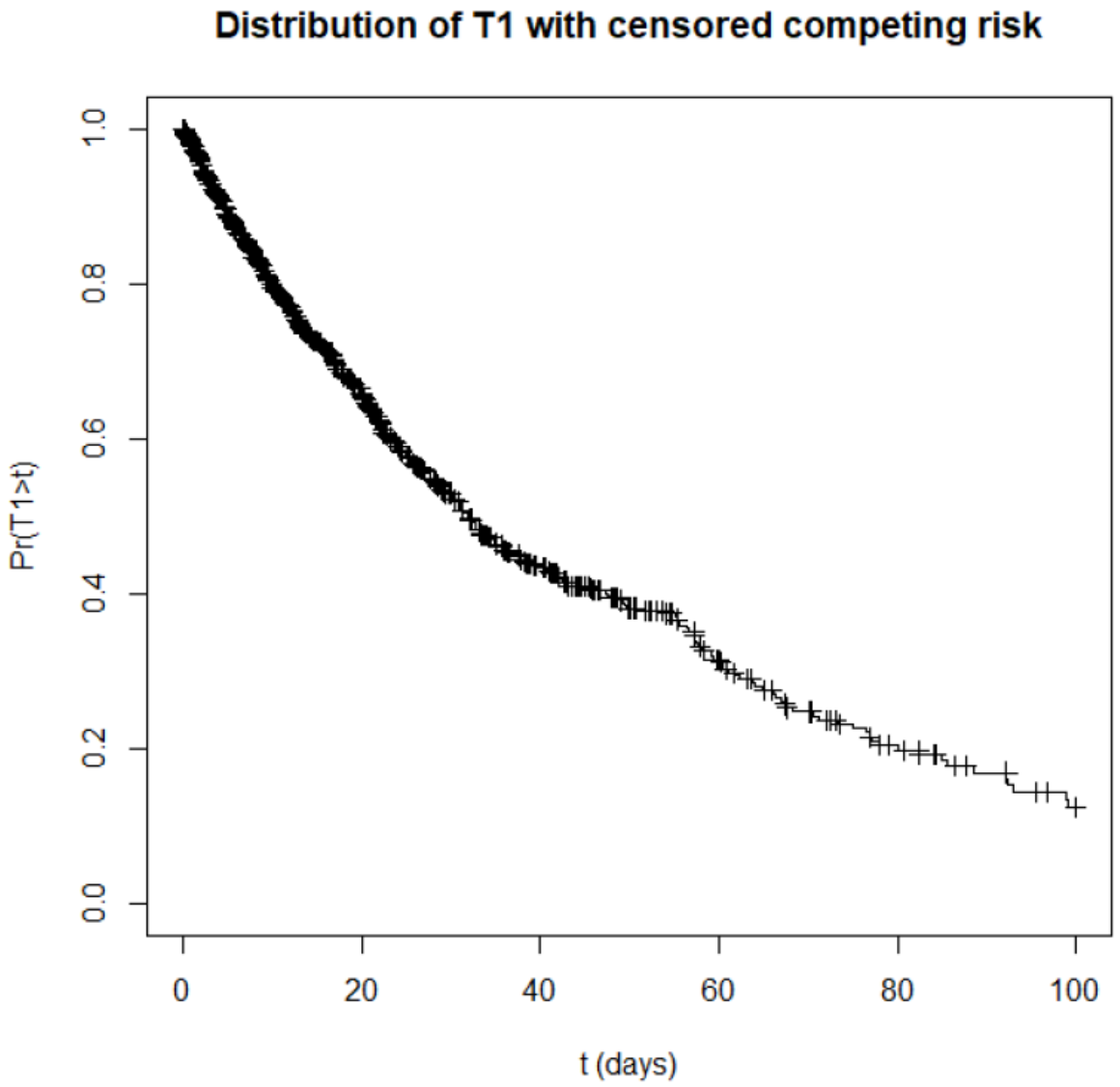
- Yet we are not able to directly observe this distribution because some events are censored at T_2



Observed data

- We can estimate the distribution of T_1, treating T_2 as censoring times

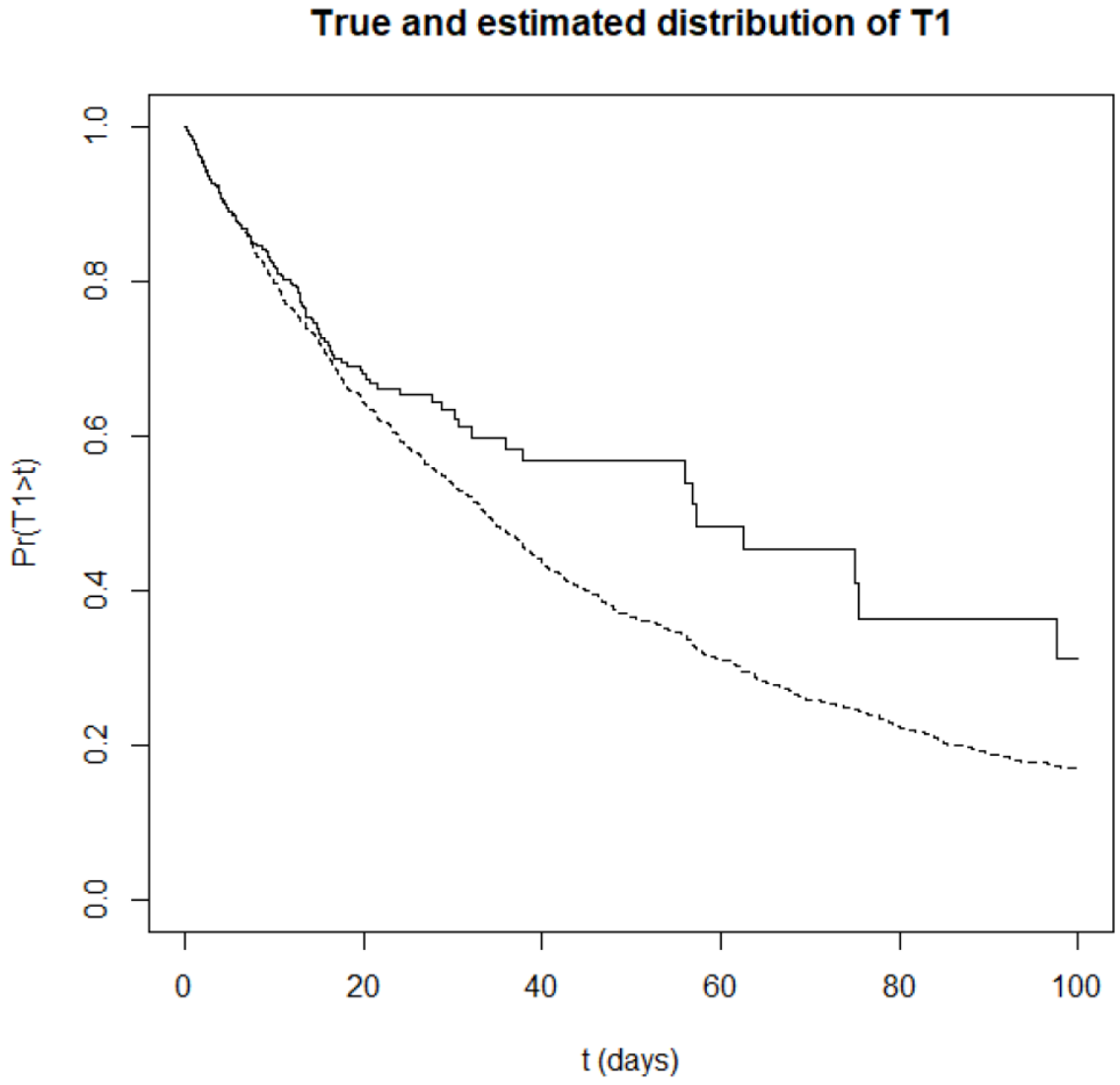
- This provides good inference when the two times are independent



Observed data

- But when the times are **dependent** (correlated), inference on T_1 can be misleading
  - *True = dotted*
  - *Observed = solid*

- Because of dependence between the event types, we cannot estimate the **type-specific survival probability** without additional assumptions



Cumulative incidence functions

- A popular alternative is to estimate the **cumulative incidence (subdistribution) function** for each event type

- This is the marginal probability of each competing event

$$F_j(t) = \Pr(T \leq t, \delta = j)$$

Cumulative incidence functions

- These summarize the probability of experiencing a given event by time t, *acknowledging that competing events may occur first*

$$F_j(\infty) < 1$$

- The cumulative incidence functions add up to the probability of experiencing *any event* by t

$$F(t) = \sum_{j=1}^{m} F_j(t)$$

VOLUME 25 · NUMBER 31 · NOVEMBER 1 2007

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Overall Survival and Cause-Specific Mortality of Patients With Stage T1a,bN0M0 Breast Carcinoma

Emer O. Hanrahan, Ana M. Gonzalez-Angulo, Sharon H. Giordano, Roman Rouzier, Kristine R. Broglio, Gabriel N. Hortobagyi, and Vicente Valero

Goal: Researchers sought to investigate the impact of prognostic factors on **breast-cancer specific (BCSM) and non-breast cancer-related mortality** in patients diagnosed with stage 1 breast cancer with a maximum tumor diameter of 1 cm.

Population: The researchers used the United States Surveillance, Epidemiology, and End Results (**SEER) program**, which covers approximately 26% of the US population. They identified patients with their target breast cancer diagnosis who were diagnosed between January 1, 1988, and December 31, 2001. This included 51,246 patients.

Hanrahan et al. (2007) JCO https://ascopubs.org/doi/10.1200/JCO.2006.08.0499

Outcome variable: The outcome of interest was time from diagnosis until death (due to any cause), date of last follow-up, or December 21, 2001. Deaths were further classified as **due to breast cancer or another cause**. Deaths from other causes were considered a competing risk in further analyses.

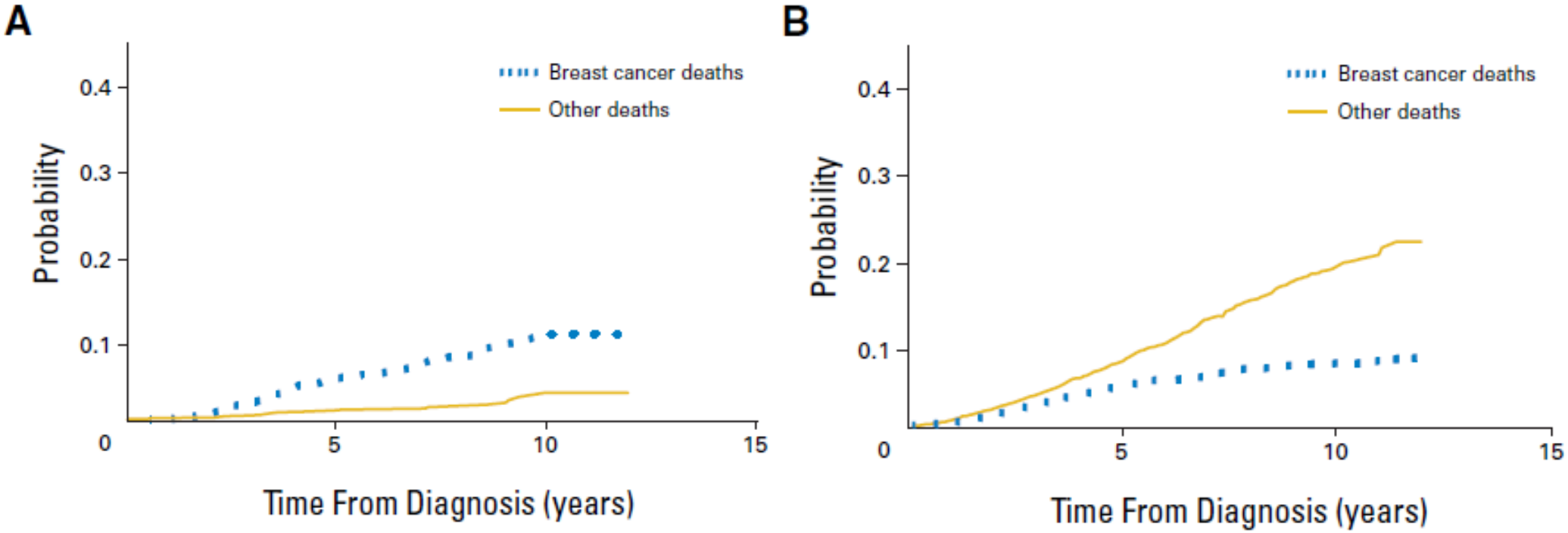
Predictor variables: The researchers considered race, laterality of the cancer (left, right), tumor grade, radiation, tumor size, histology group, breast surgery type, lymph node surgery type, estrogen receptor (ER) positivity, and progesterone receptor (PR) positivity. ER and PR status were recorded only for patients diagnosed after 1990 and were associated with a large number of missing values.

Statistical analysis: Researchers fit a Cox proportional hazards model to overall survival.

For their competing risks analysis, they estimated the probability of death resulting from breast cancer along with the probability of death resulting from other causes.

Analyses were stratified on age (<50, ≥50) and ER status.

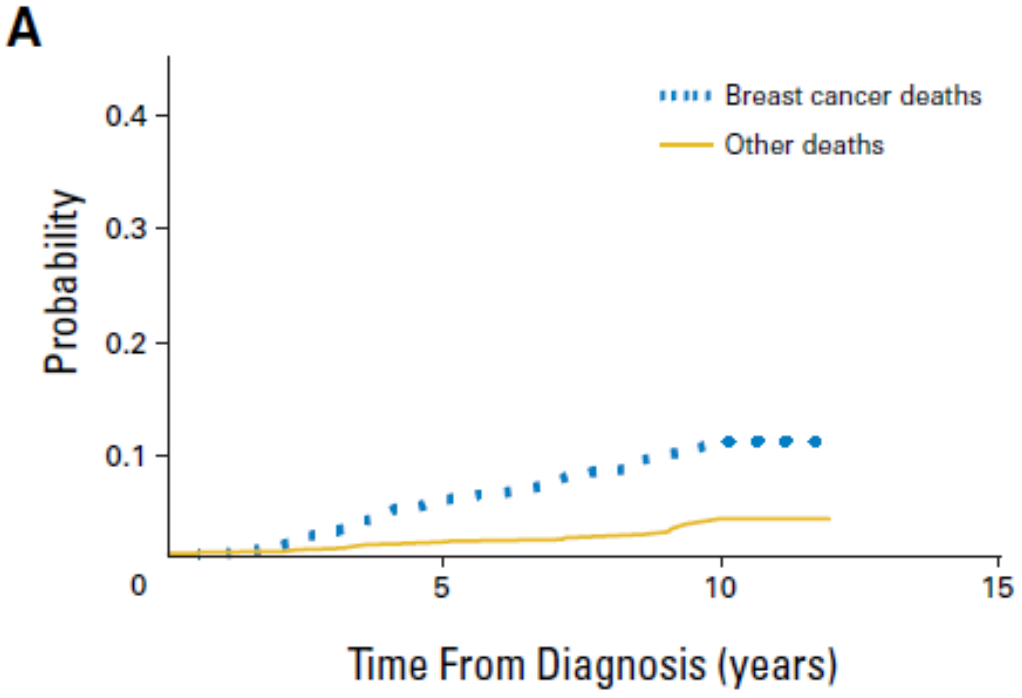
Results: The figure below summarizes the probability of death from breast cancer and death from other causes by age and ER status. Panel (A) is for women age <50 years and ER negative. Panel (B) is for women age ≥50 and ER negative.



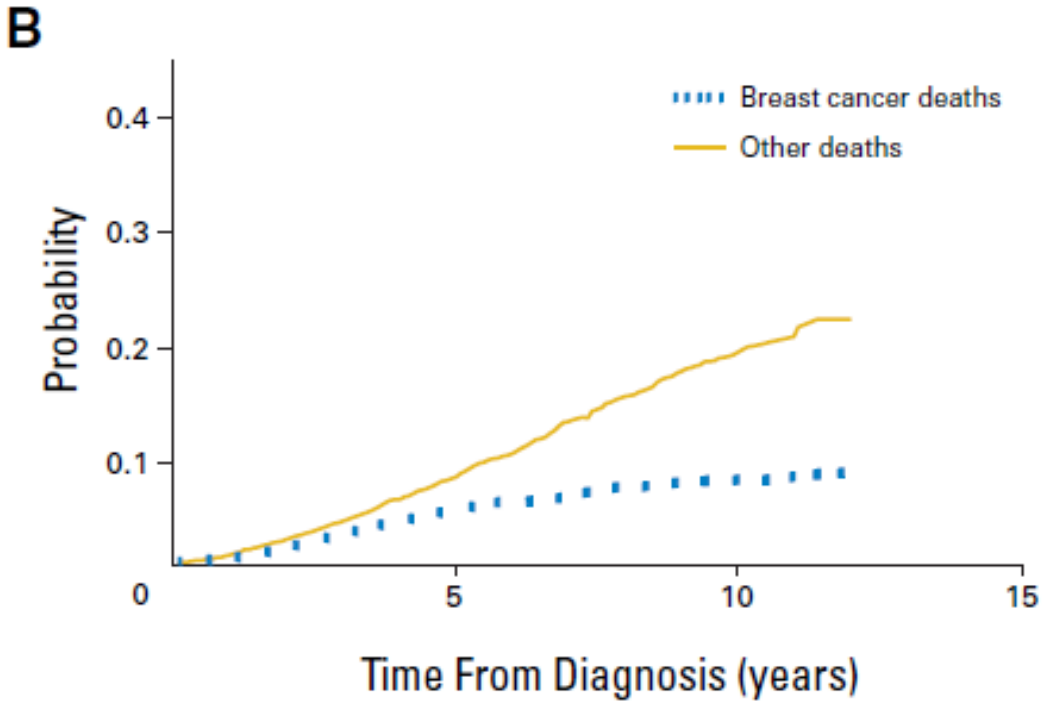
Younger women

Older women

We see that older women (Panel B) have higher overall mortality than younger women. This can be inferred by adding the blue and yellow curves together. This difference in overall mortality is driven by the probability of non-cancer related deaths, which are more frequent in older women, as expected.

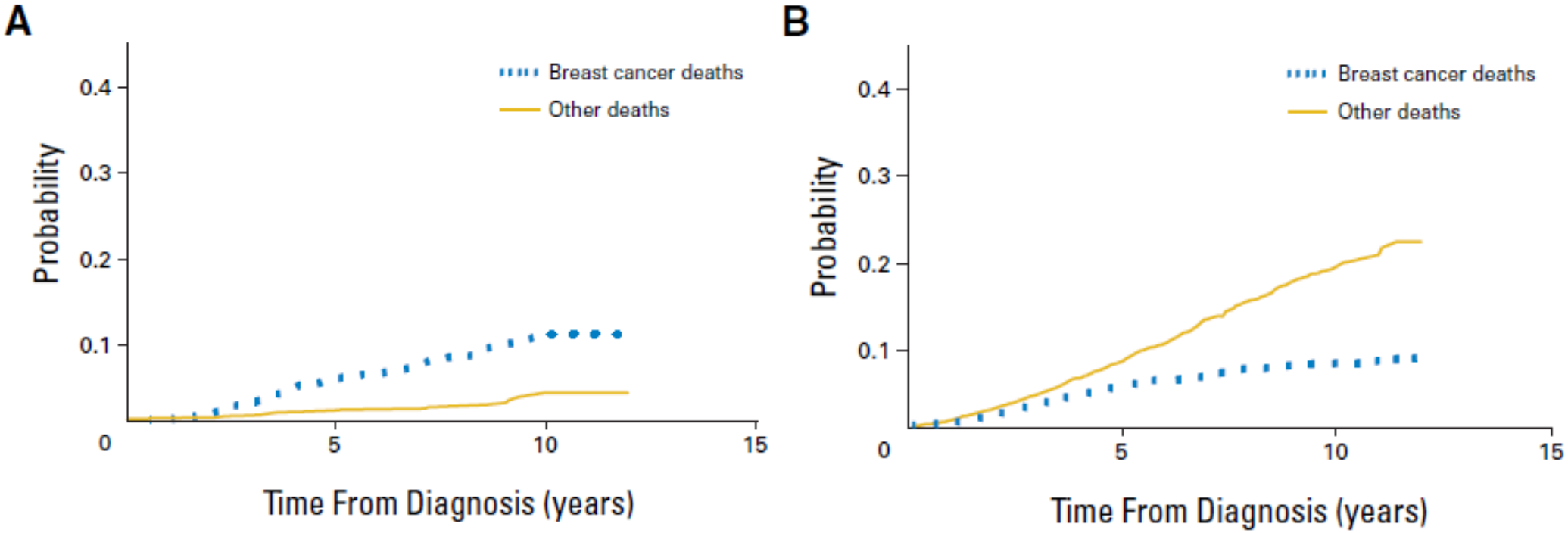


Younger women



Older women

Importantly, breast cancer-related deaths represent a higher proportion of overall deaths in younger women (Panel A). Women under 50 diagnosed with this cancer subtype are more likely to die from breast cancer than other causes, which in turn impacts the risk/benefit profile of certain treatment strategies. This important feature is not captured in an analysis of all-cause mortality.



Younger women

Older women

ORIGINAL ARTICLE

Lifetime Risks of Cardiovascular Disease

Jarett D. Berry, M.D., Alan Dyer, Ph.D., Xuan Cai, M.S., Daniel B. Garside, B.S., Hongyan Ning, M.D., Avis Thomas, M.S., Philip Greenland, M.D., Linda Van Horn, R.D., Ph.D., Russell P. Tracy, Ph.D., and Donald M. Lloyd-Jones, M.D.

<u>Goal</u>: Estimate **lifetime risks of cardiovascular disease (CVD)** across the age spectrum in black adults and white adults

<u>Population</u>: Meta-analysis using data from 18 cohort studies involving a total of **257,384 adults**. Each study included 10 or more years of follow-up.

<u>Outcome variable</u>: The primary endpoint was **CVD death**. Death by other causes was treated as a **competing risk**.

The time scale was age, and the **time origin** was age 55.

<u>Predictor variables</u>: Sex, race (black or white), and birth cohort.

Health-related risk factors: diabetes, current smoking, cholesterol, and systolic blood pressure.

Individuals were placed into **risk factor groups**:

(1) all risk factors within the optimal range, (2) ≥ 1 risk factor not optimal, (3) ≥ 1 risk factor elevated, (4) 1 major risk factor, or (5) ≥ 2 major risk factors.

Risk factor group modeled as a **time-dependent covariate**.

<u>Statistical analysis</u>: **Competing risks** framework to calculate the cumulative incidence of CVD death for each age attained during follow-up.

Participant data were **stratified according to sex and risk-factor group**. Secondary analyses considered stratifying by race and birth cohort.

Authors note that Kaplan-Meier approach "*overestimates the remaining lifetime risk of cardiovascular disease when the competing risk is high*."

<u>Results</u>: **Figure 1** summarizes lifetime risk of death from cardiovascular disease among black and white men at 55 years of age, stratified by risk factor group.

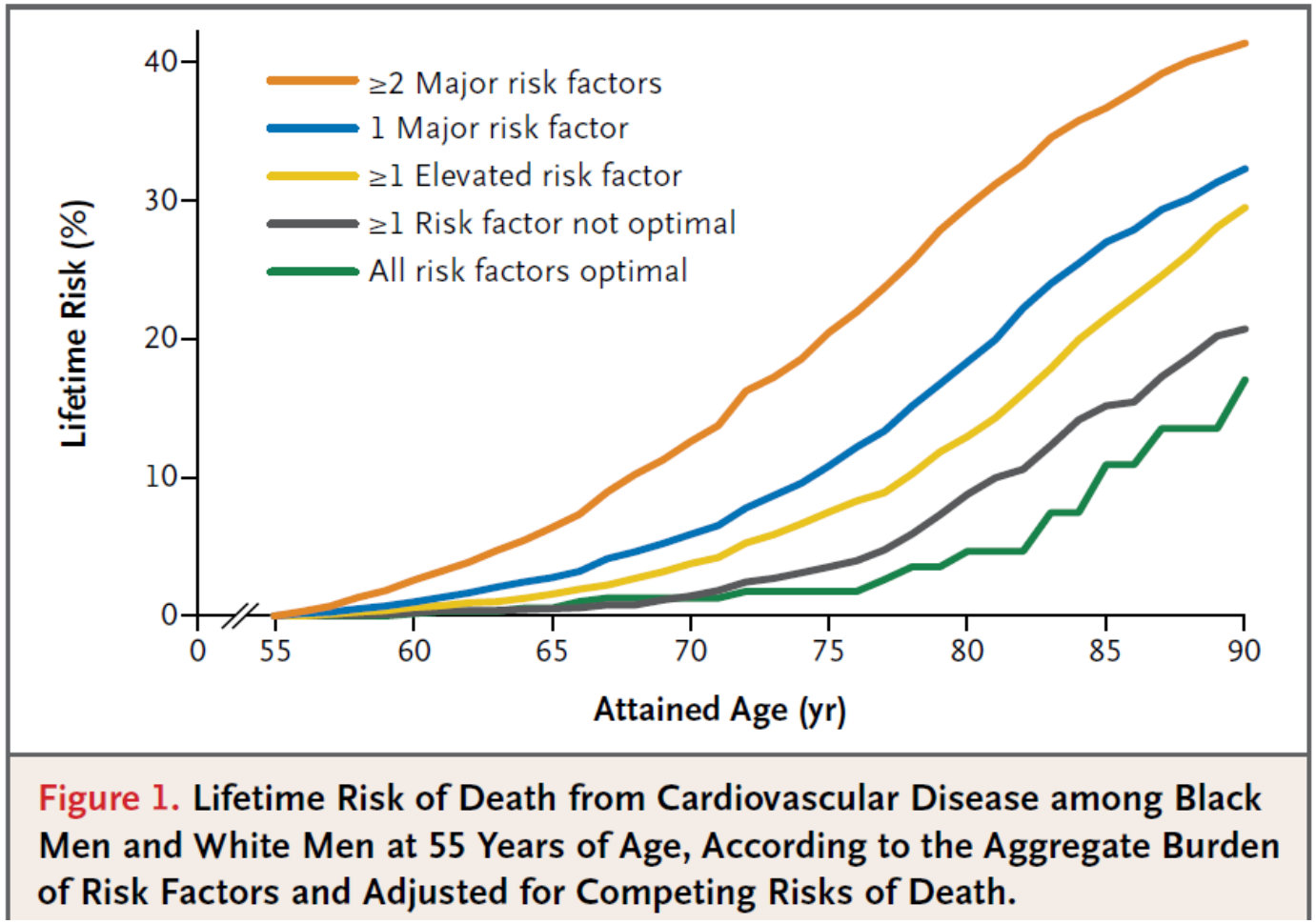
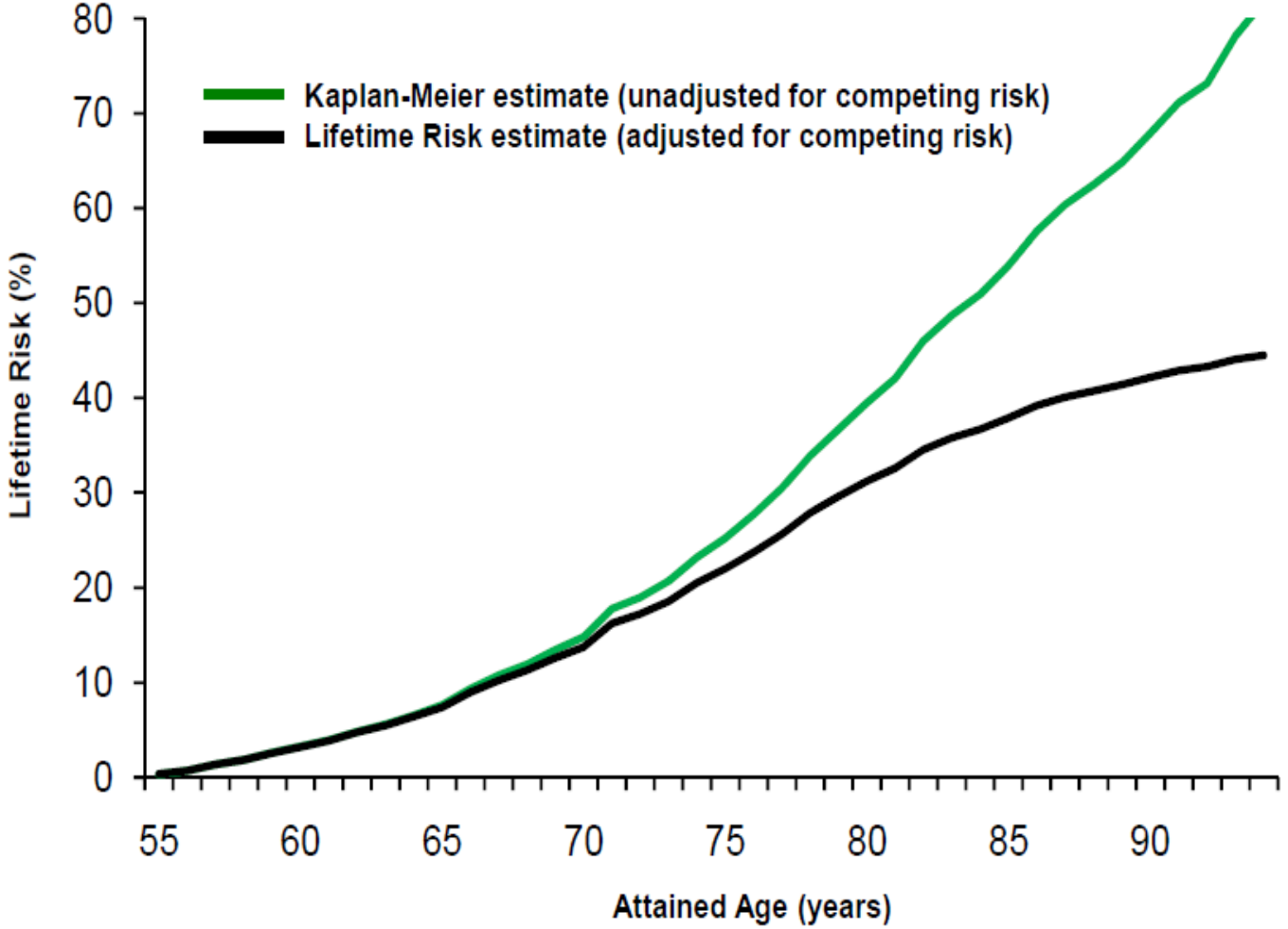


Figure 1. Lifetime Risk of Death from Cardiovascular Disease among Black Men and White Men at 55 Years of Age, According to the Aggregate Burden of Risk Factors and Adjusted for Competing Risks of Death.

Lifetime risk low among persons who had an optimal risk-factor profile at all ages

Corresponding cumulative incidence functions for non-CVD death not shown

Supplemental **Figure 5a** compares cumulative incidence in highest risk group (black) with naïve Kaplan-Meier estimate (green)



Kaplan-Meier overestimates lifetime risk

Individuals who have died from other causes are not at risk for future CVD death

Conclusion

- I hope you have found this material useful for your studies
- I greatly value your feedback on the course content

- I have enjoyed our time together this semester
- Please stay in touch!

nataliedean@emory.edu