Advanced Epidemiologic Methods

EPID 722
Spring 2021
UNC – Chapel Hill
jessedwards@unc.edu

Instructors

Jess Edwards

Steve Cole

TAs

Linnea Olson

Rachael Ross

What are we doing here?

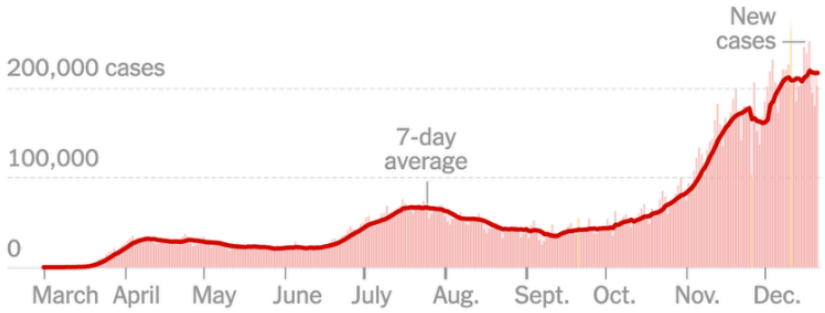
One goal of epidemiology is to describe the world as it is.

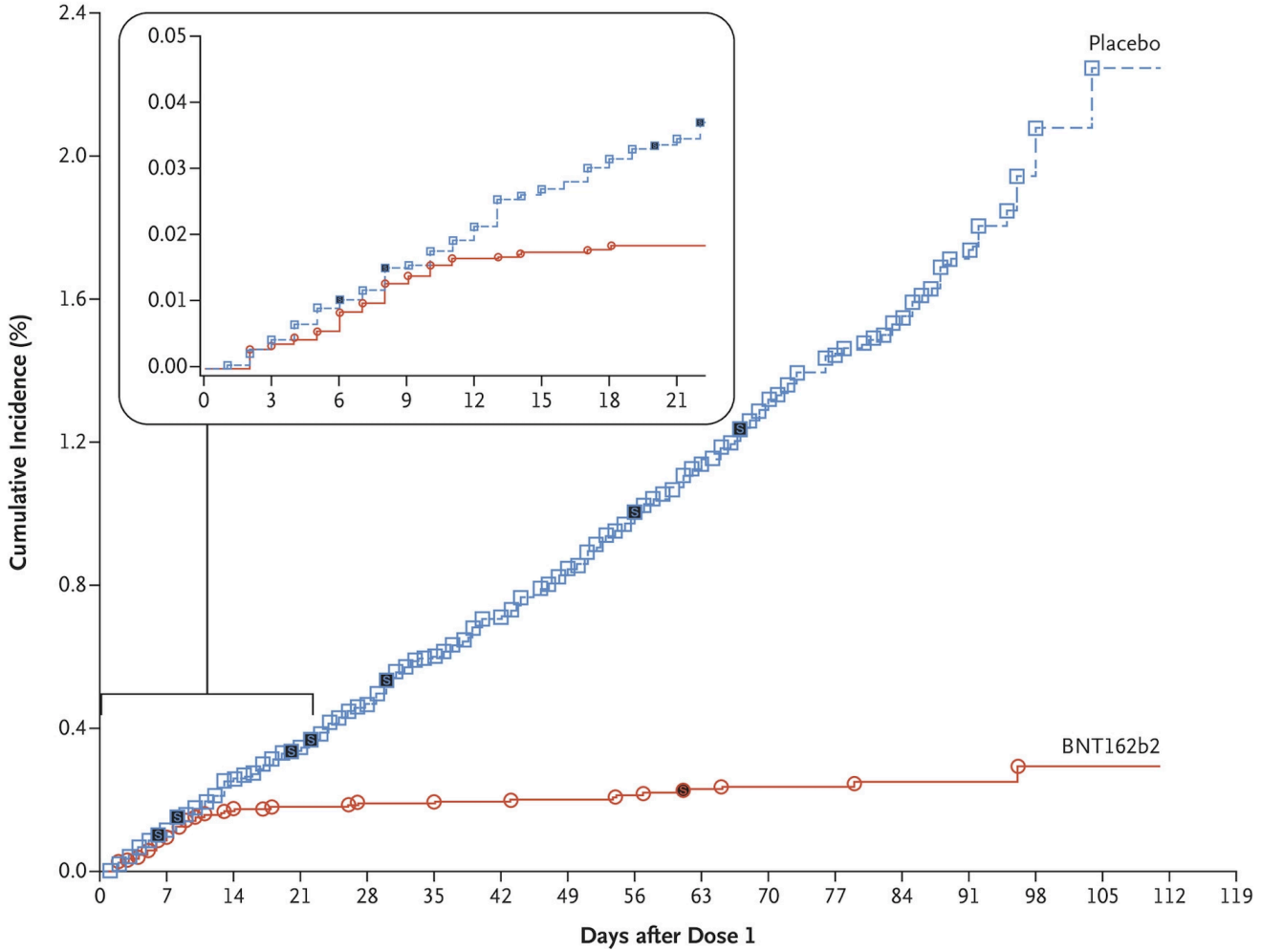
Where does disease occur? What is the burden of a specific disease or condition?
Who gets sick? Are there health disparities? Where is need greatest?
Are there associations between health status and specific substances, treatments, or behaviors?

Another goal is to describe the world as it could be under some intervention

How might health outcomes change under more rapid treatment?
Would disease be reduced by banning a specific chemical?
Would behavior modification lead to a reduction in incidence?

This is a fascinating time to be an epidemiologist.





Why train in advanced epidemiologic methods?

1. INTRODUCTION

The subject-specific data from a longitudinal study consist of a string of numbers. These numbers represent a series of empirical measurements. Calculations are performed on these strings and causal inferences are drawn. For example, an investigator might conclude that the analysis provides strong evidence for "a direct effect of AZT on the survival of AIDS patients controlling for the intermediate variable – therapy with aerosolized pentamidine". The nature of the relationship between the sentence expressing these causal conclusions and the computer calculations performed on the strings of numbers has been obscure. Since the computer algorithms are well-defined mathematical objects, it is useful to provide formal mathematical definitions for the English sentences expressing the investigator's causal inferences, In Robins (1986, 1987), I proposed a

Describing our data is not the same as describing the world.

Our data only update our beliefs about the world to the extent that we believe key assumptions.

(Data lead us towards the truth to the extent that these required assumptions are correct (enough)).

In this class, we will discuss these assumptions and methods used to relax some of them.

If we have “perfect” data, epidemiology is easy.

What makes data “perfect”?

An incomplete list:

- No measurement error
- No missing data
- Complete follow-up
- Includes a census of the target population
- No confounding
- And so on

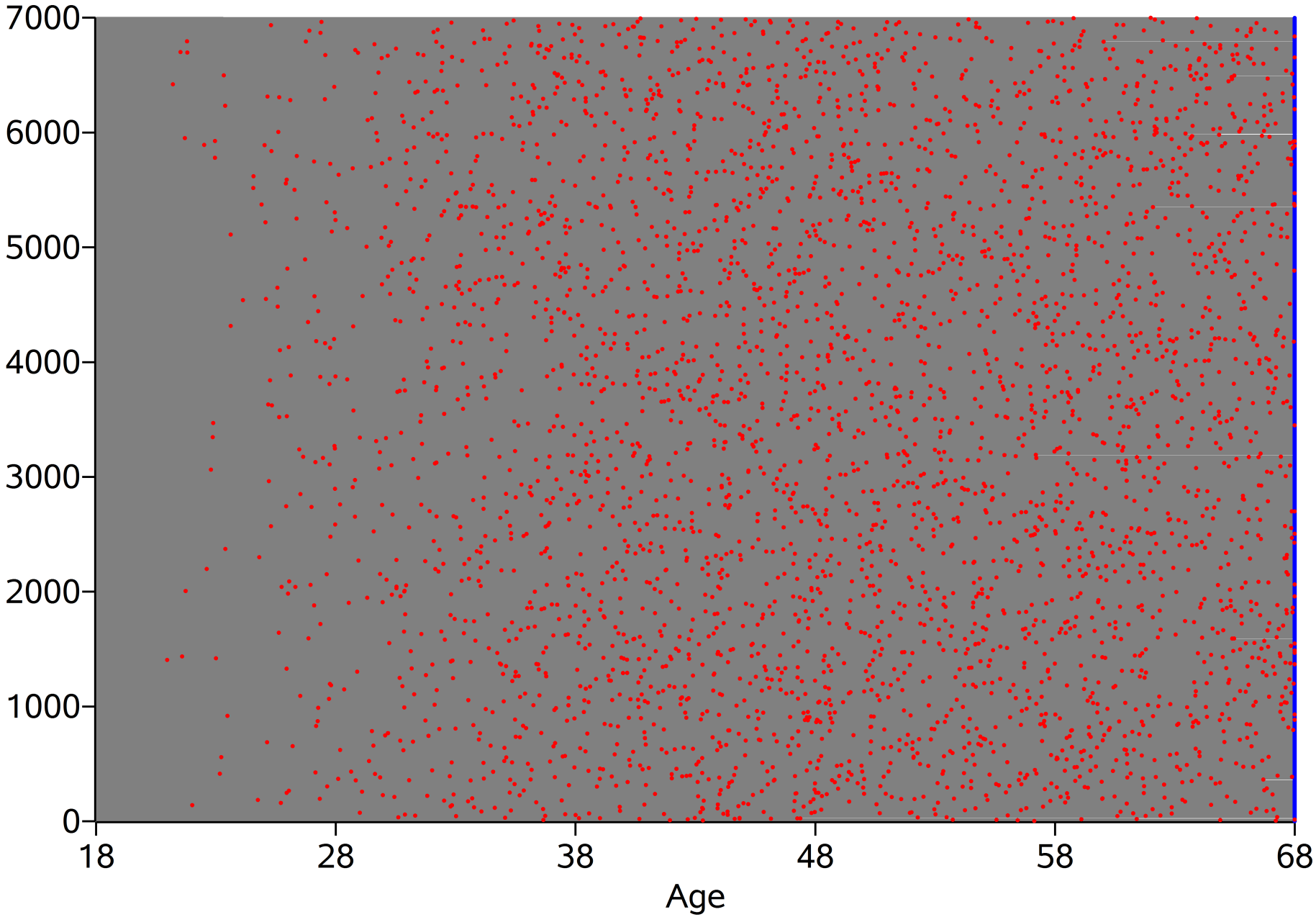
But we can’t wait for perfect data

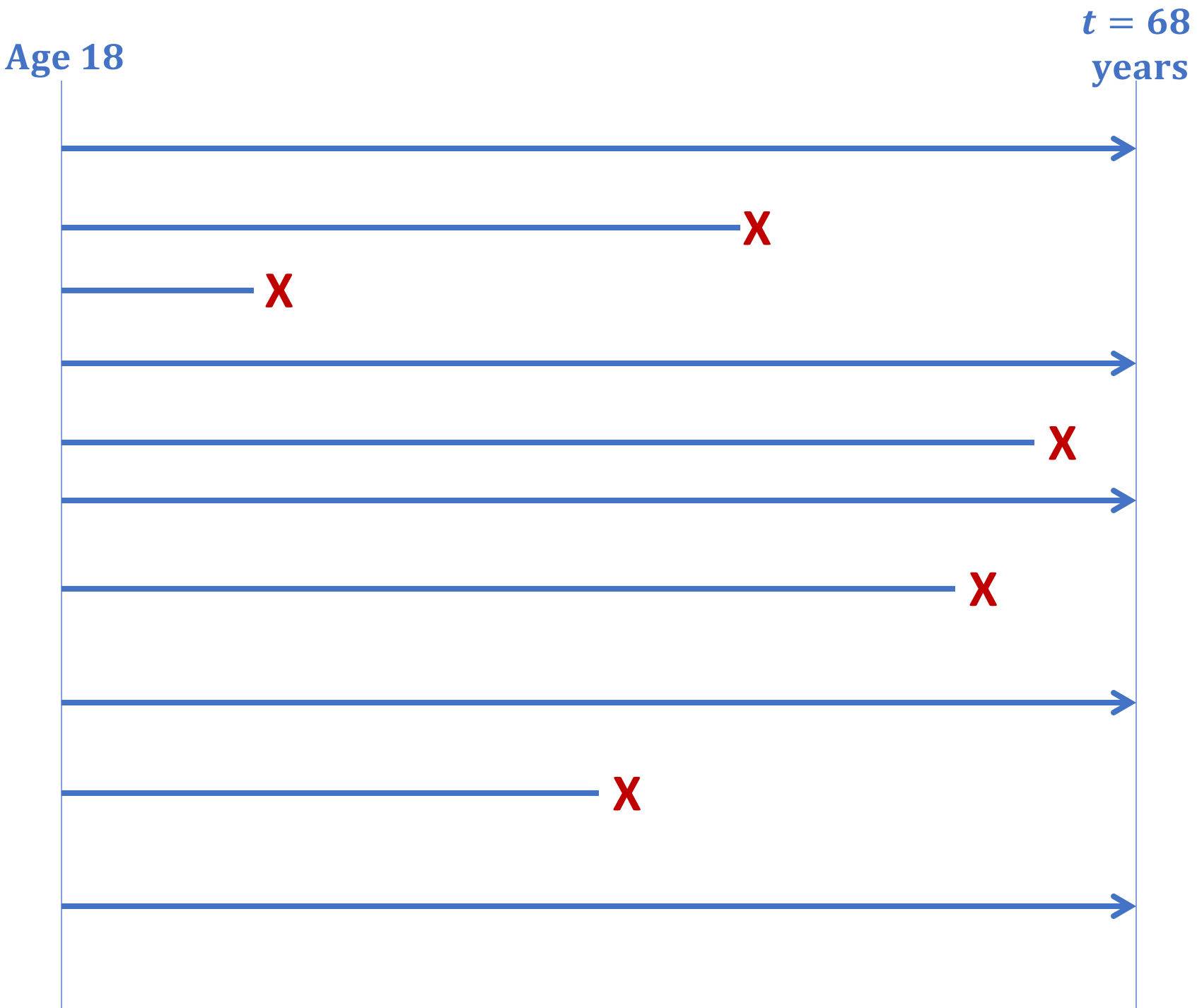
Results from epidemiologic studies power decision making.

Decisions don’t wait. (Not to decide is to decide).

So, the ability to learn from our (almost always) imperfect data is critical.

An example.

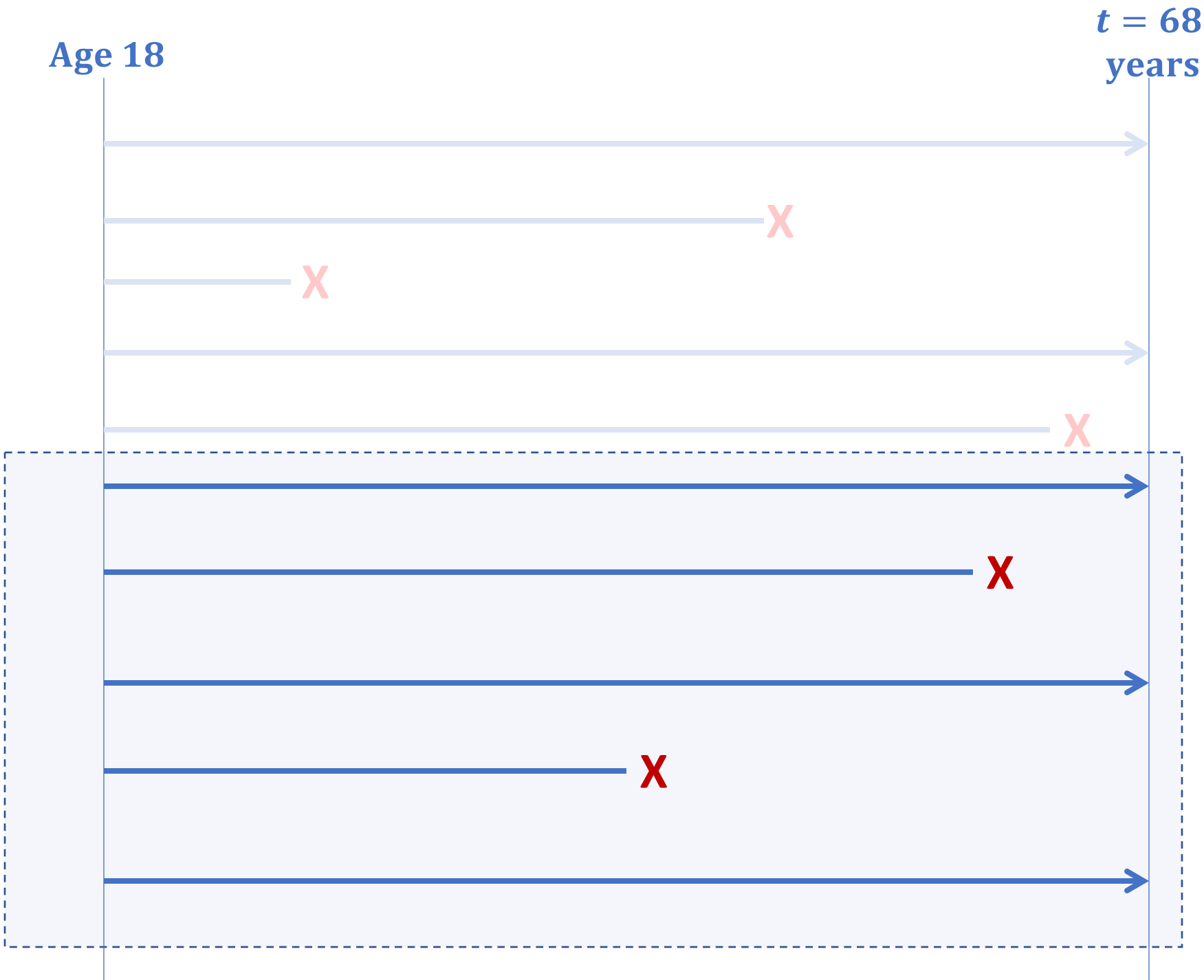


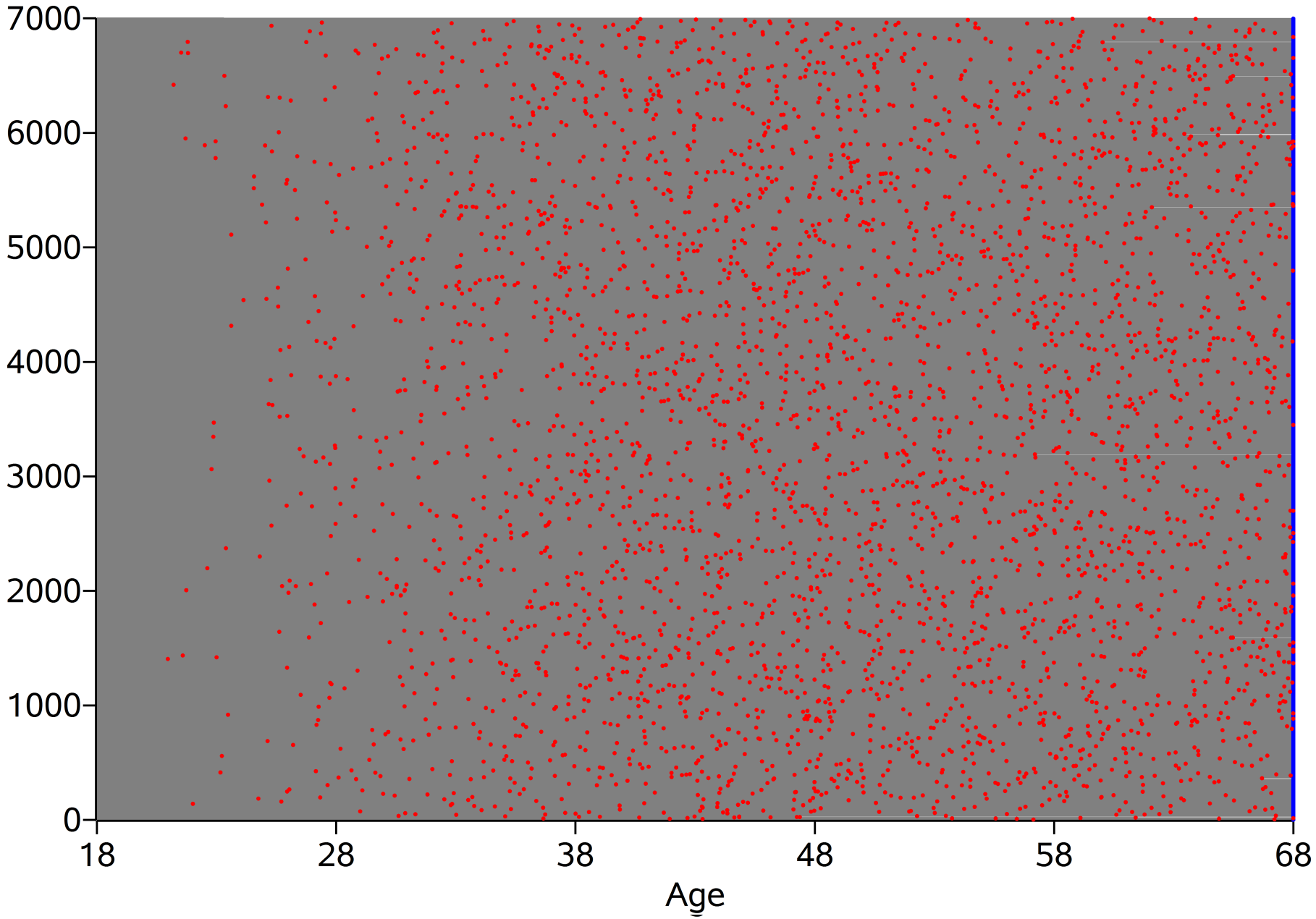

Age 18
$t = 68$ years
X
X
X
X
X

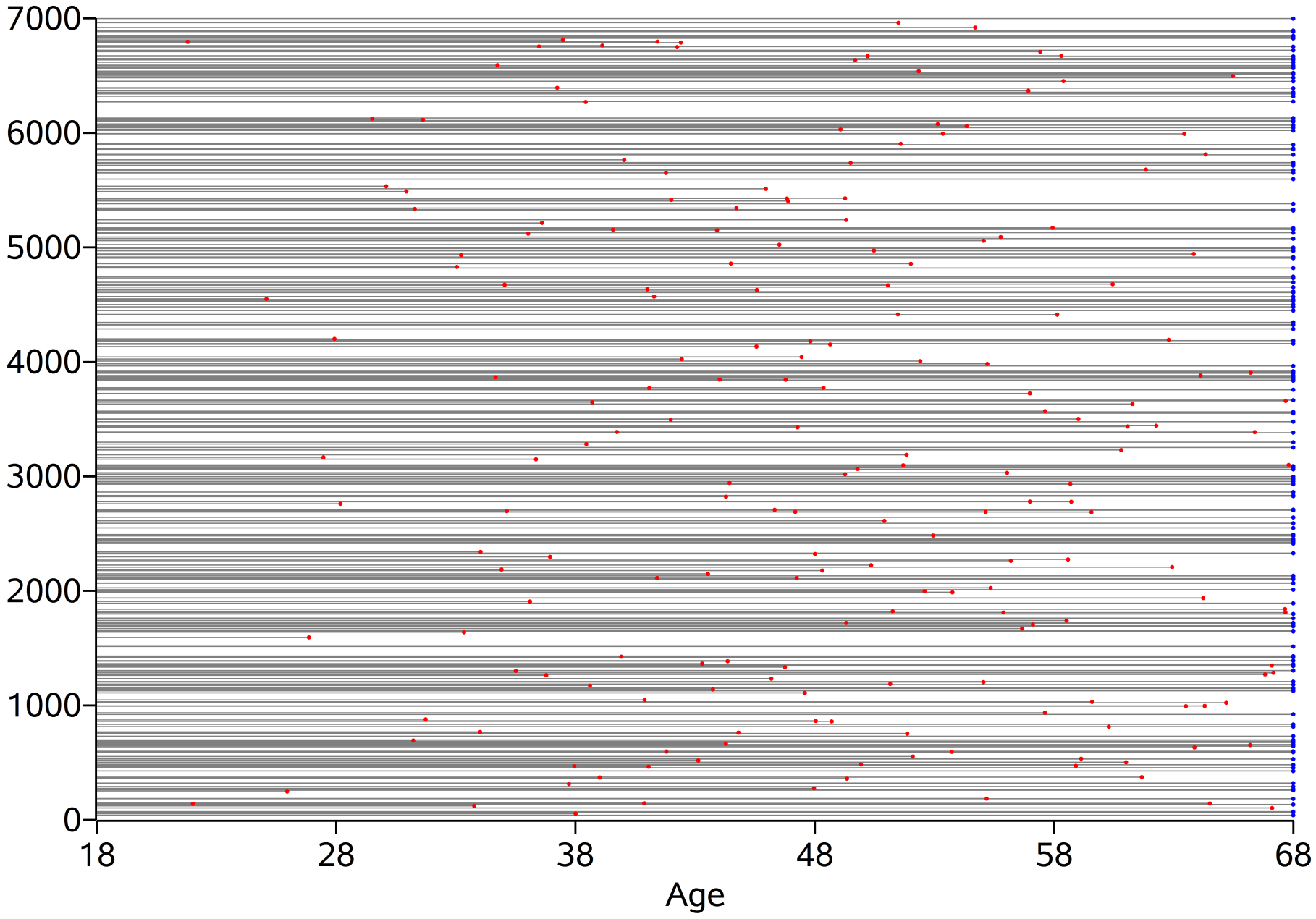
But what if we don't observe $\boldsymbol{i = 1, \ldots 7000}$?

We could fail to observe some subjects due to

- Sampling from the population
- Refusal to participate
- Selective recruitment into the study
- Other reasons?


Age 18
$t = 68$ years
X
X
X
X
X
X


7000
6000
5000
4000
3000
2000
1000
0
18
28
38
48
58
68
Age

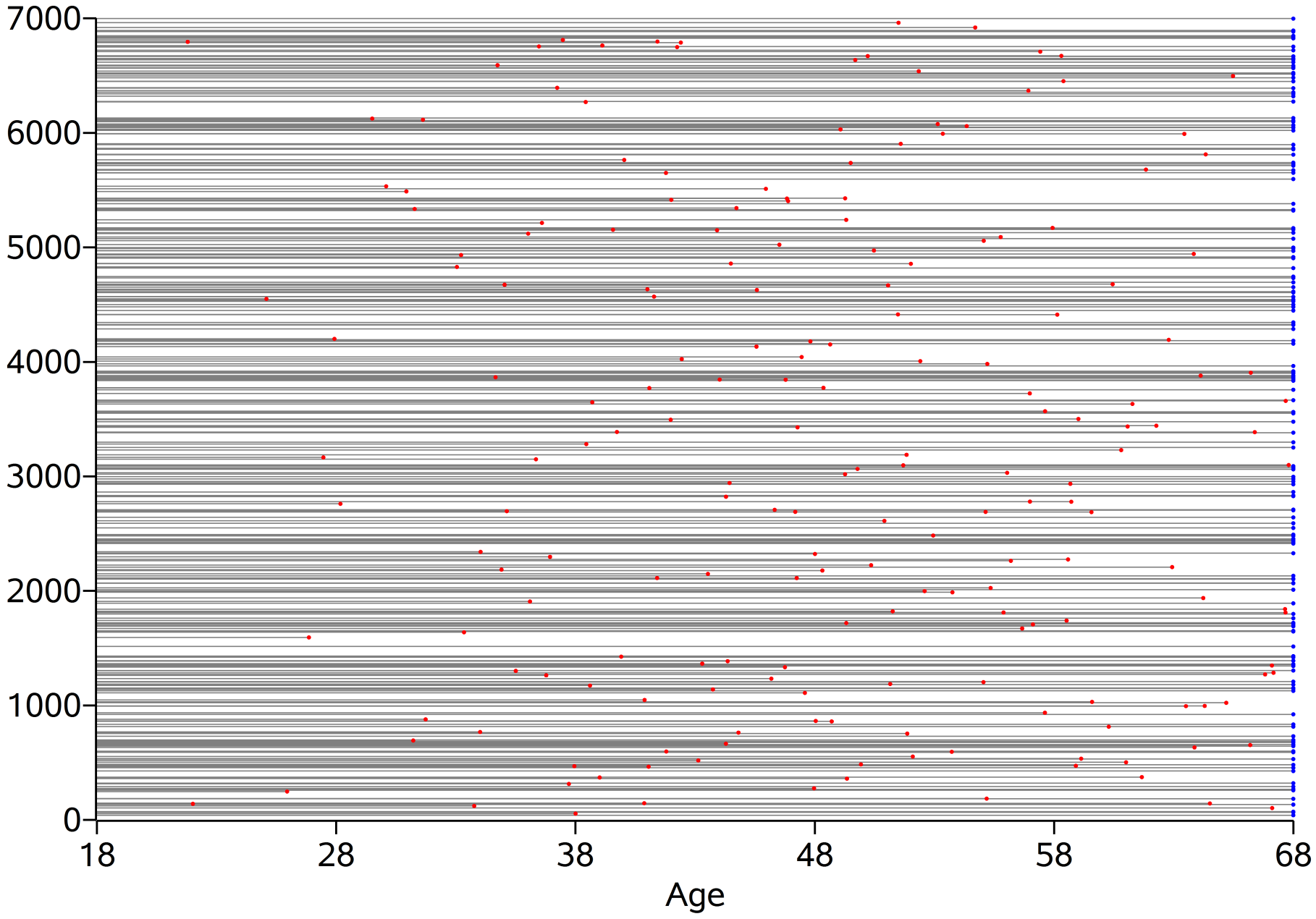

7000
6000
5000
4000
3000
2000
1000
0
18
28
38
48
58
68
Age

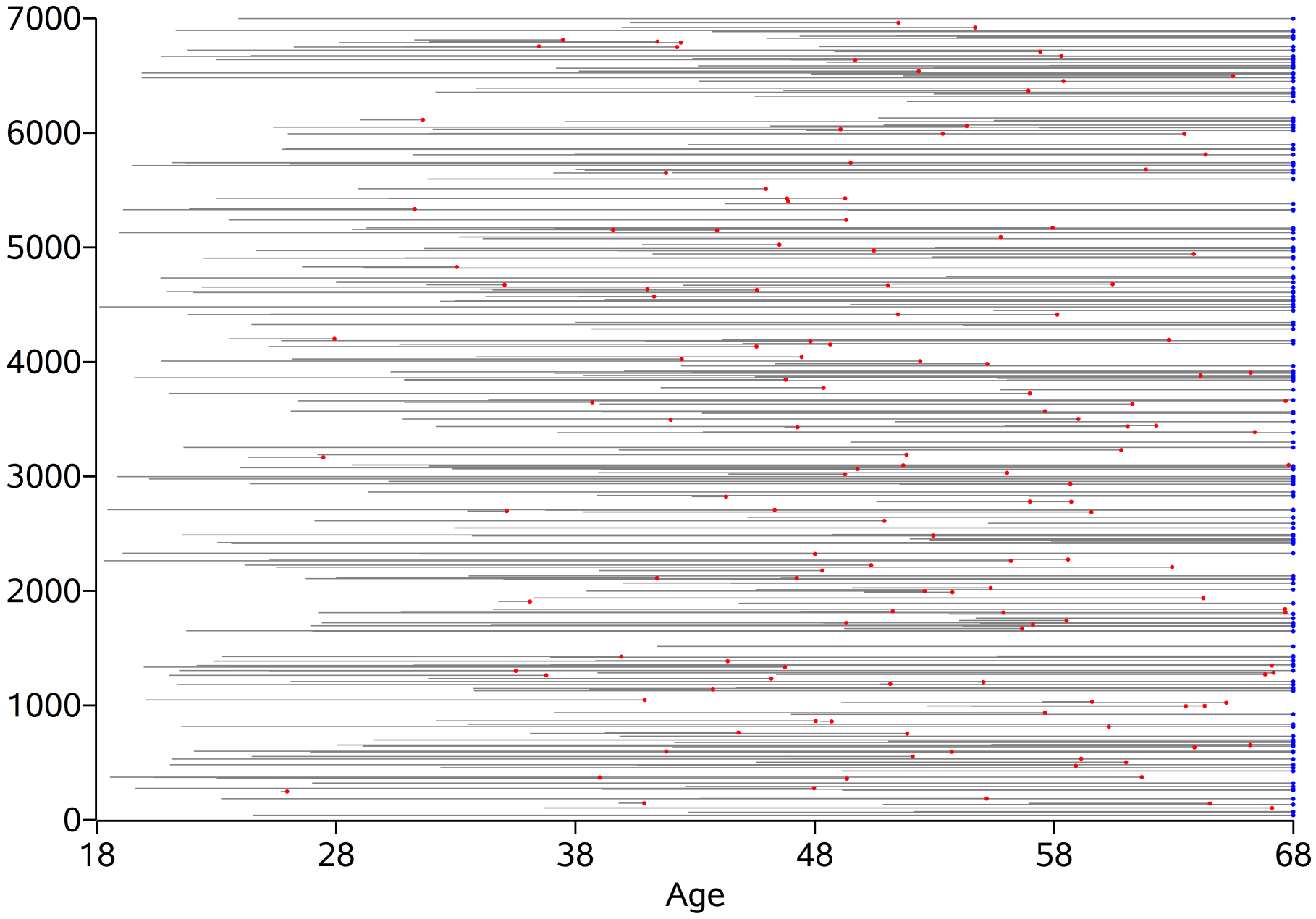
What if some participants are not recruited into the study until after age 18?

We may enroll subjects into the study late due to

- Delayed identification of eligible subjects
- Migration
- Others?


Age 18
$t = 68$ years
X
X
X
X
X


7000
6000
5000
4000
3000
2000
1000
0
18
28
38
48
58
68
Age

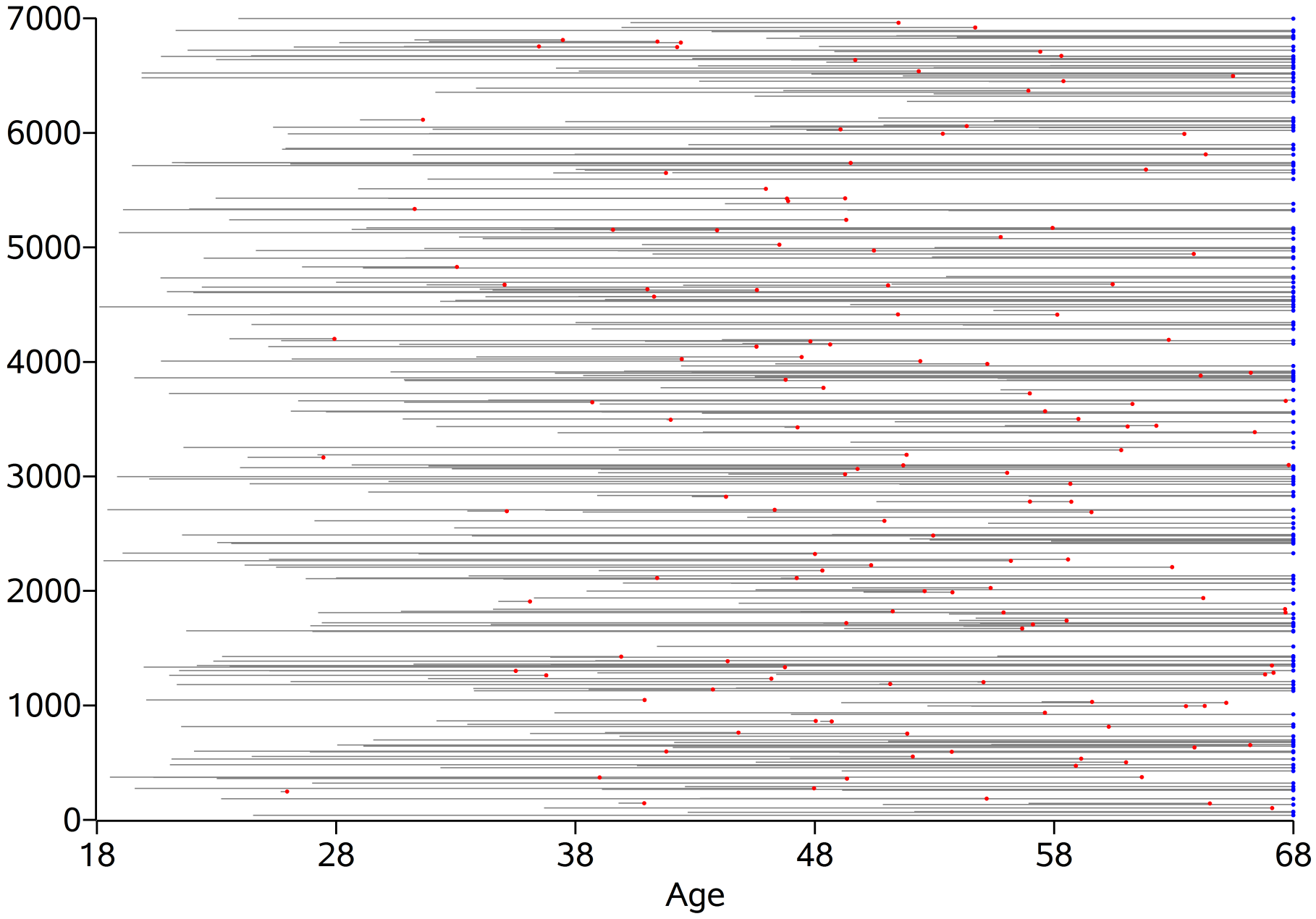

7000
6000
5000
4000
3000
2000
1000
0
18
28
38
48
58
68
Age

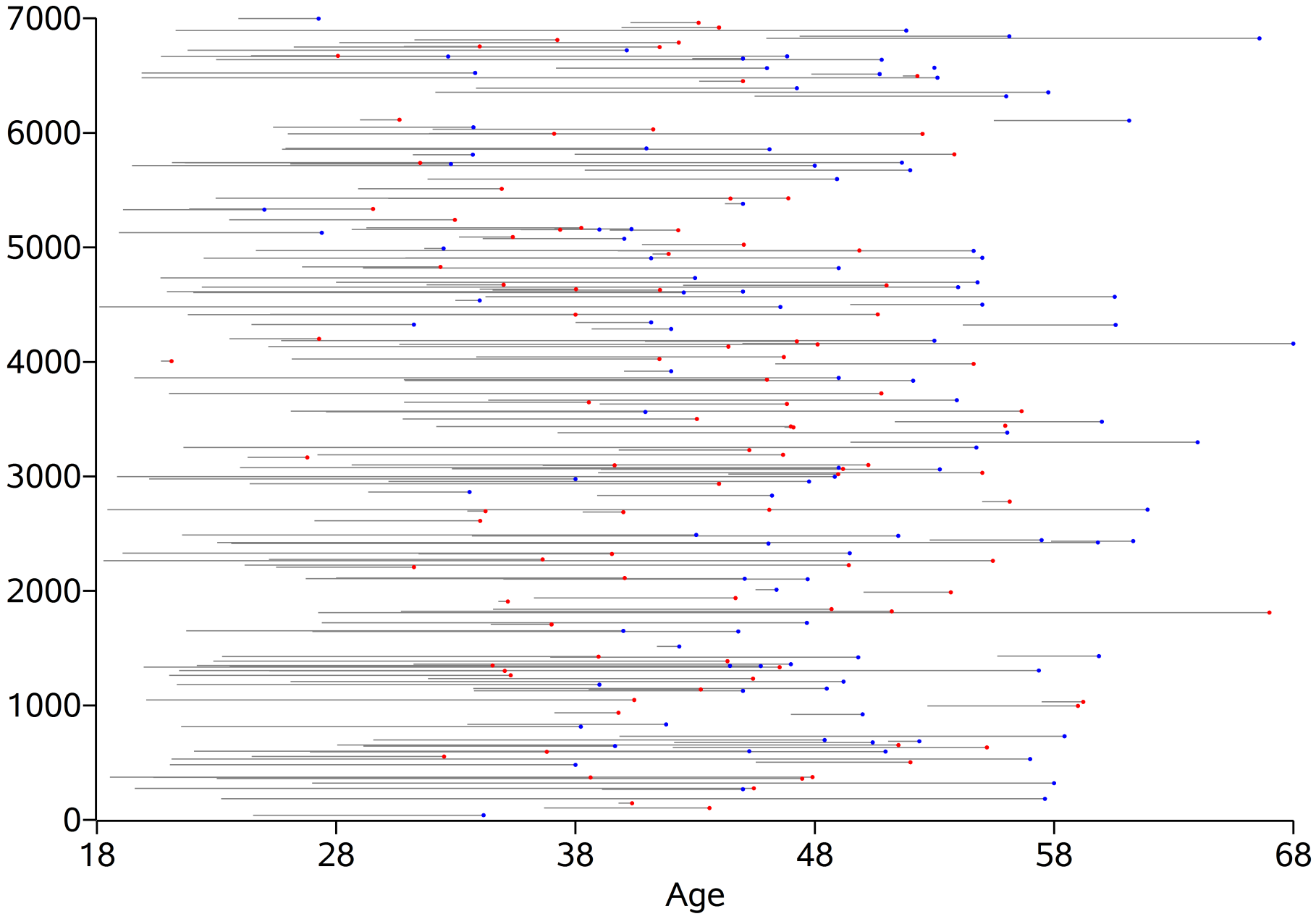
Now, what if we don’t observe the all of the selected participants until age 68?

We could fail to observe some subjects for the full study period due to

- Loss to follow-up
- Administrative end of the study in 2017
- Others?


Birth
$t = 68$ years
A
0
1
X
0
X
1
0
X
1
1
X
0
0
X
1

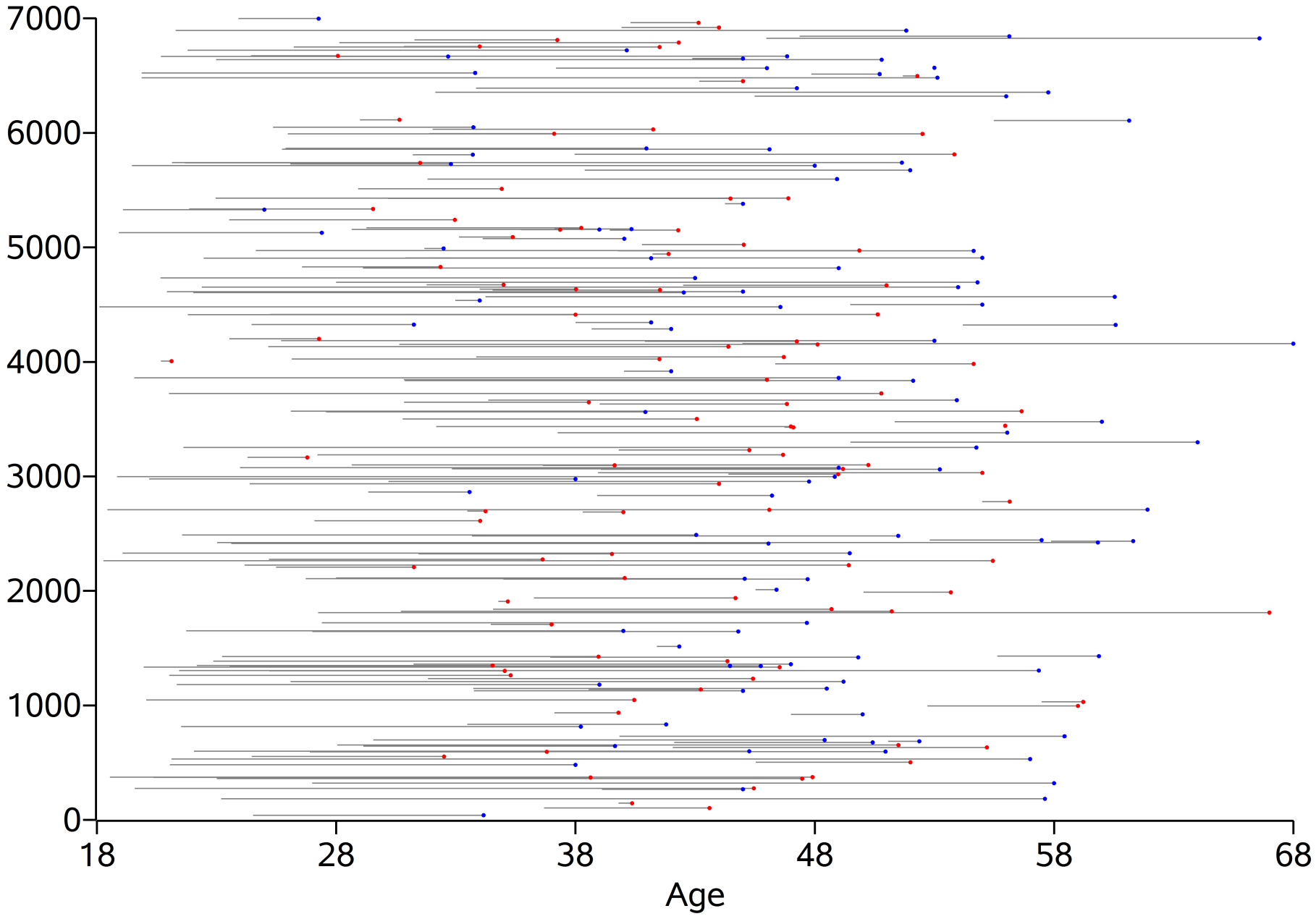

7000
6000
5000
4000
3000
2000
1000
0
18
28
38
48
58
68
Age

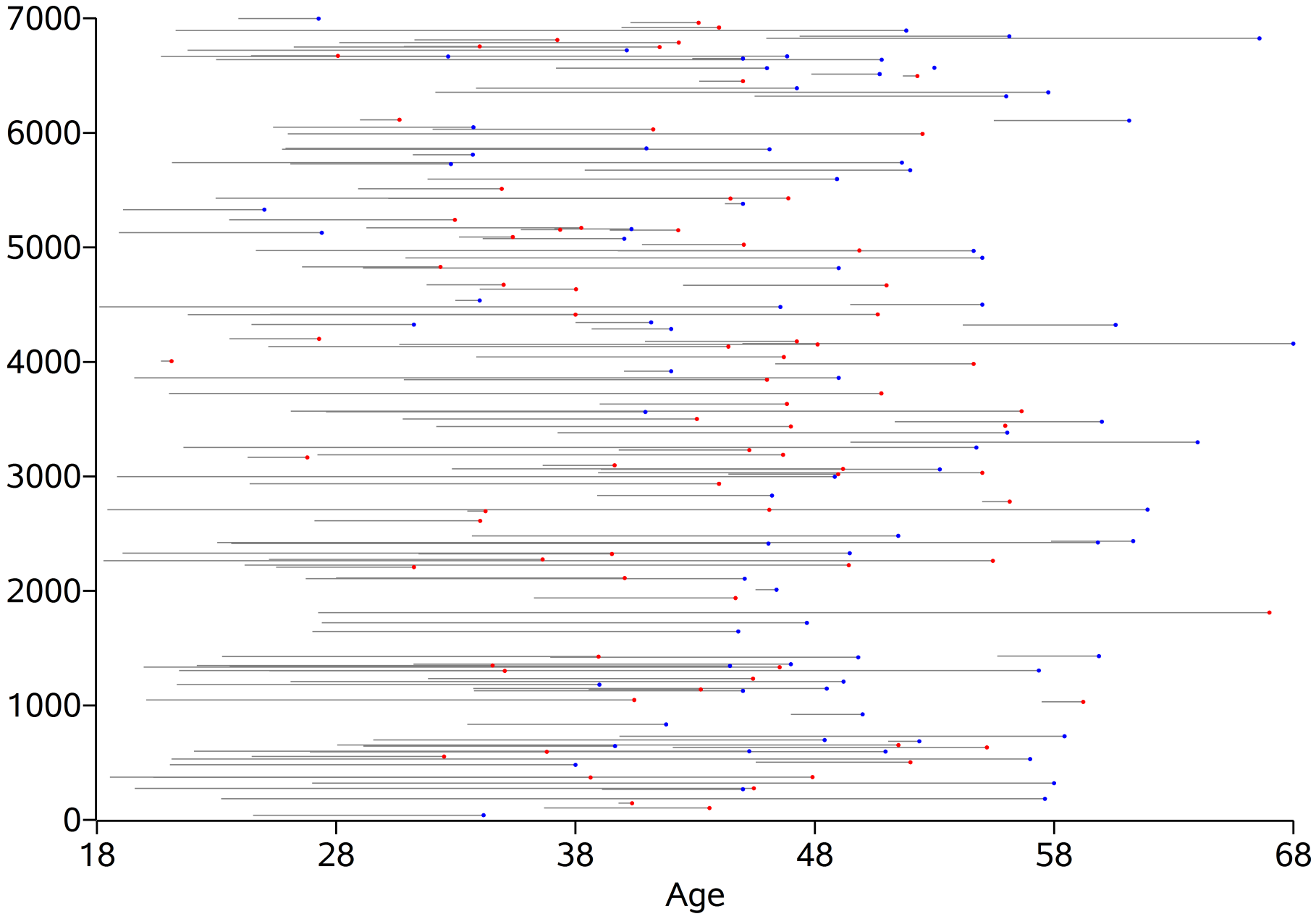

7000
6000
5000
4000
3000
2000
1000
0
18
28
38
48
58
68
Age

And subjects in the study may have missing values for important variables

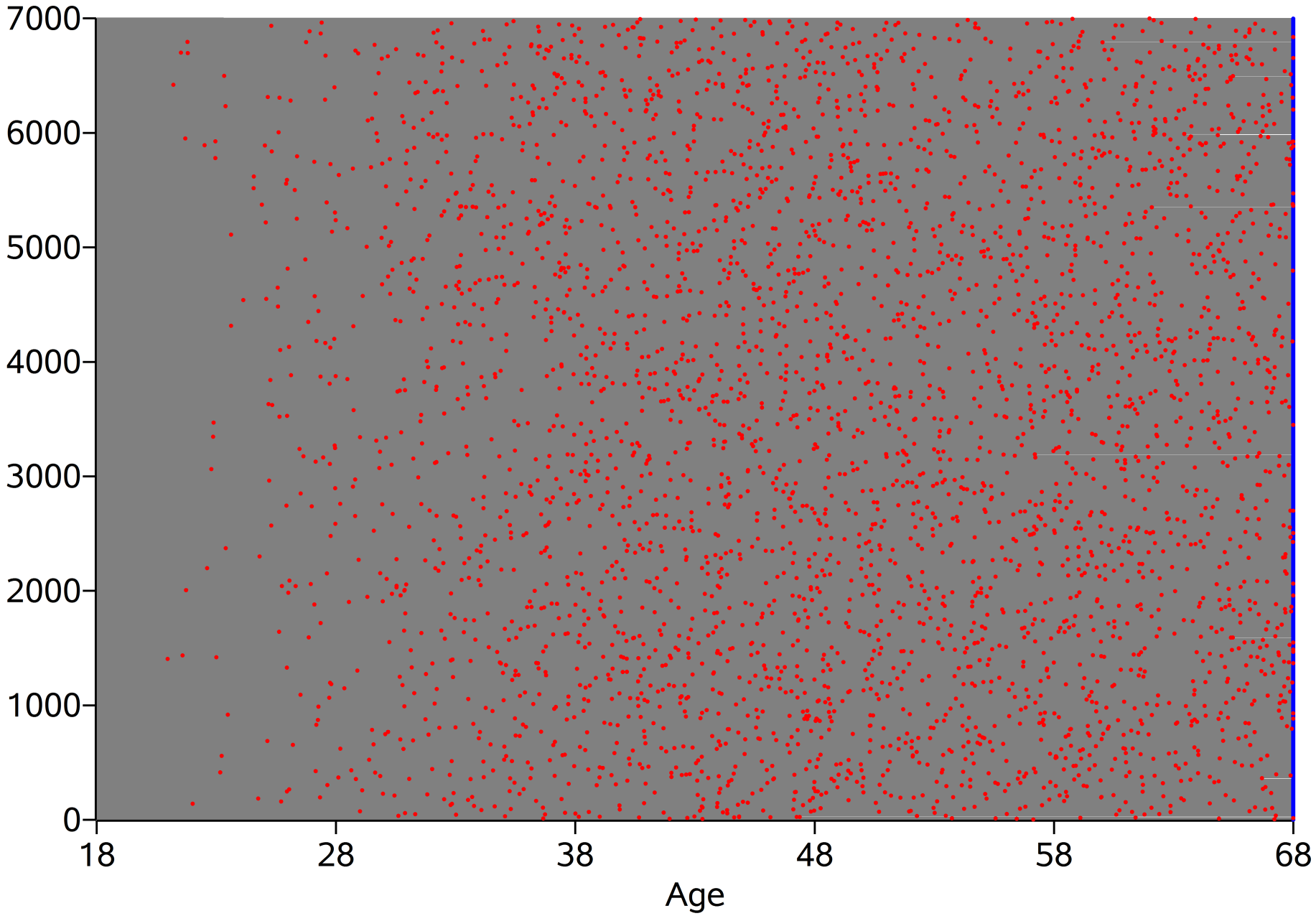
Subjects may have missing values for important variables due to

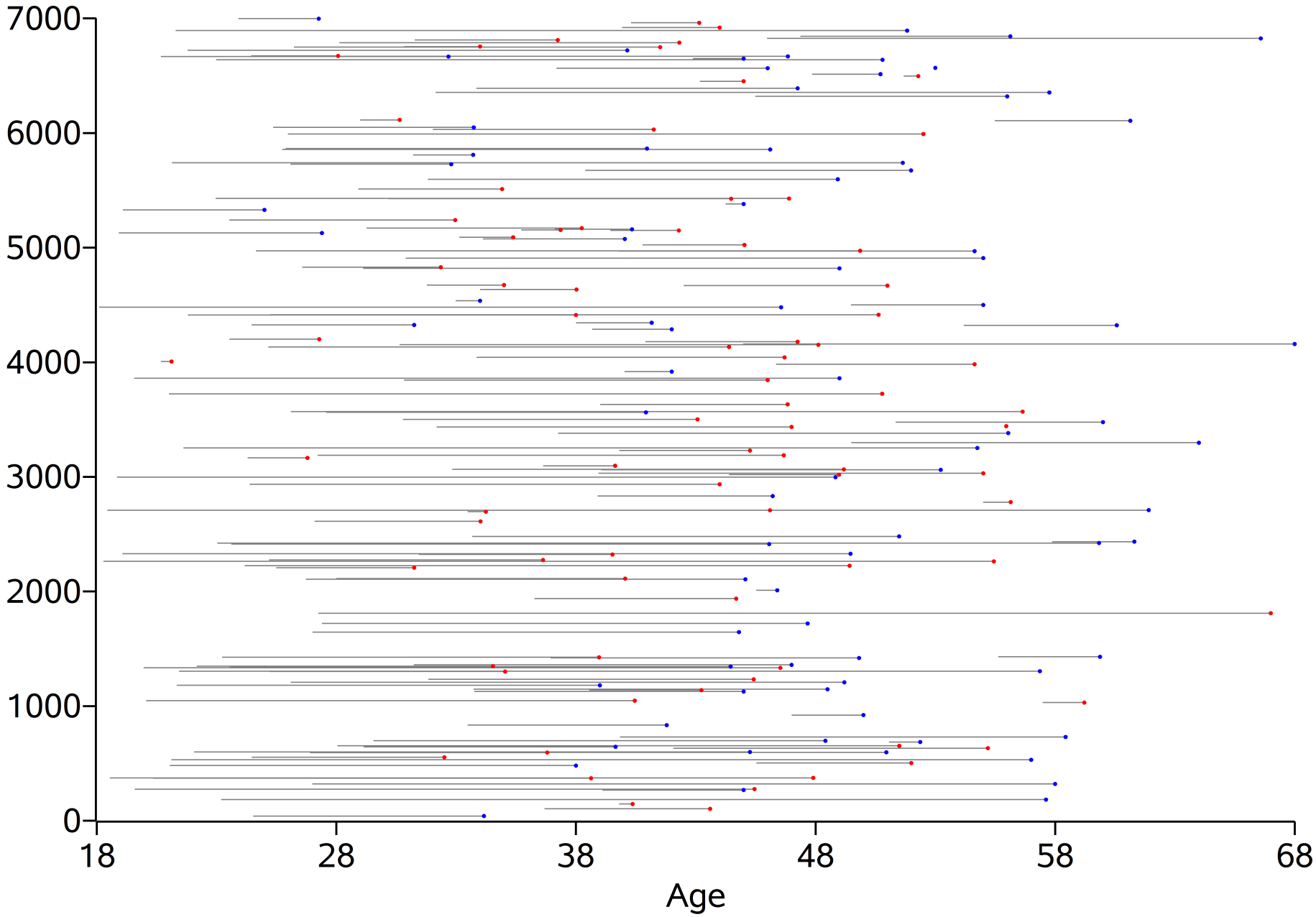
- Errors in data entry
- Subject does not know requested information (e.g., vaccinations)
- Subject refuses to provide information (e.g., drug use)
- Assay failures
- Assay detection limits


7000
6000
5000
4000
3000
2000
1000
0
18
28
38
48
58
68
Age


7000
6000
5000
4000
3000
2000
1000
0
18
28
38
48
58
68
Age

to reiterate the amount of missing information…..


7000
6000
5000
4000
3000
2000
1000
0
18
28
38
48
58
68
Age


7000
6000
5000
4000
3000
2000
1000
0
18
28
38
48
58
68
Age

Course Roadmap

A course in 3 acts

Asking and framing epidemiologic questions

Describing the world as it is: Tools for survival analysis

Describing the world under interventions: Tools for causal inference

Asking and framing questions

What are the components of a good question in epidemiology?

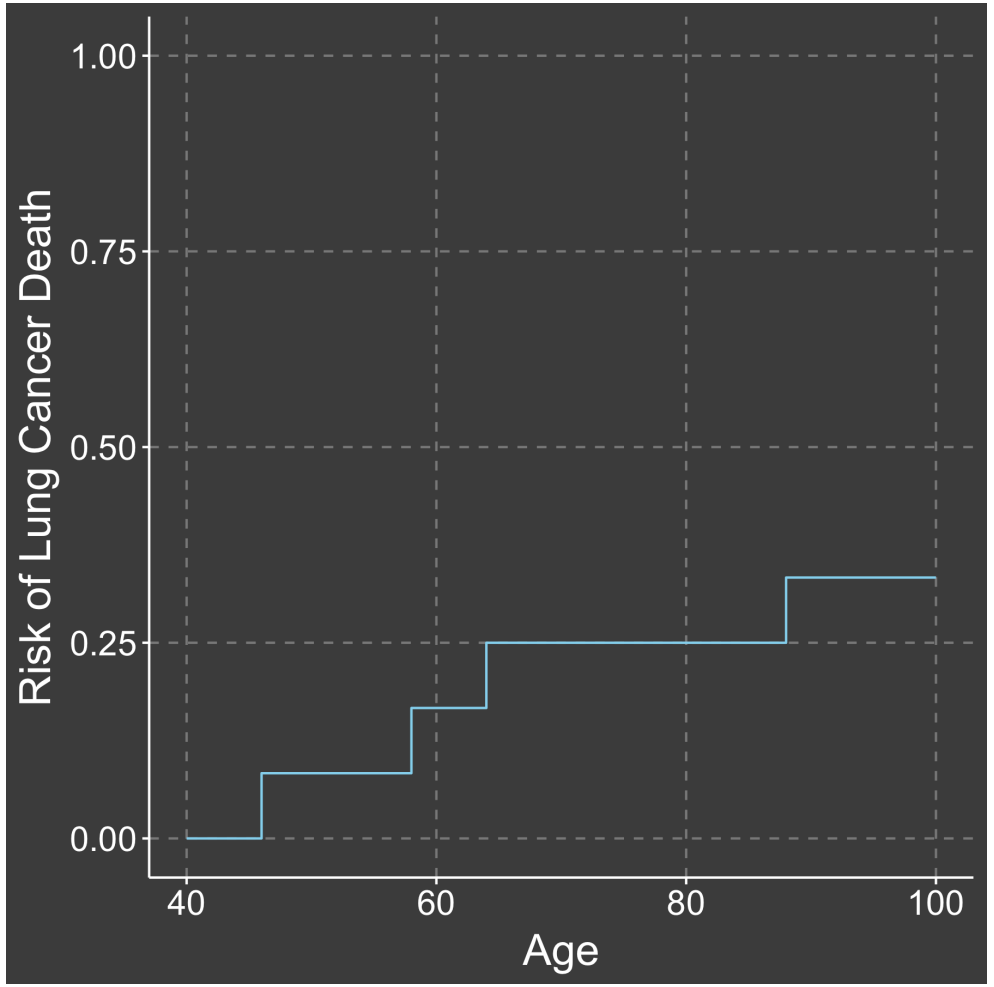
Describing the world as it is

We will focus on estimating risk.

(Why risk?)

We will start with the “single sample” scenario.

(Why 1 sample?)

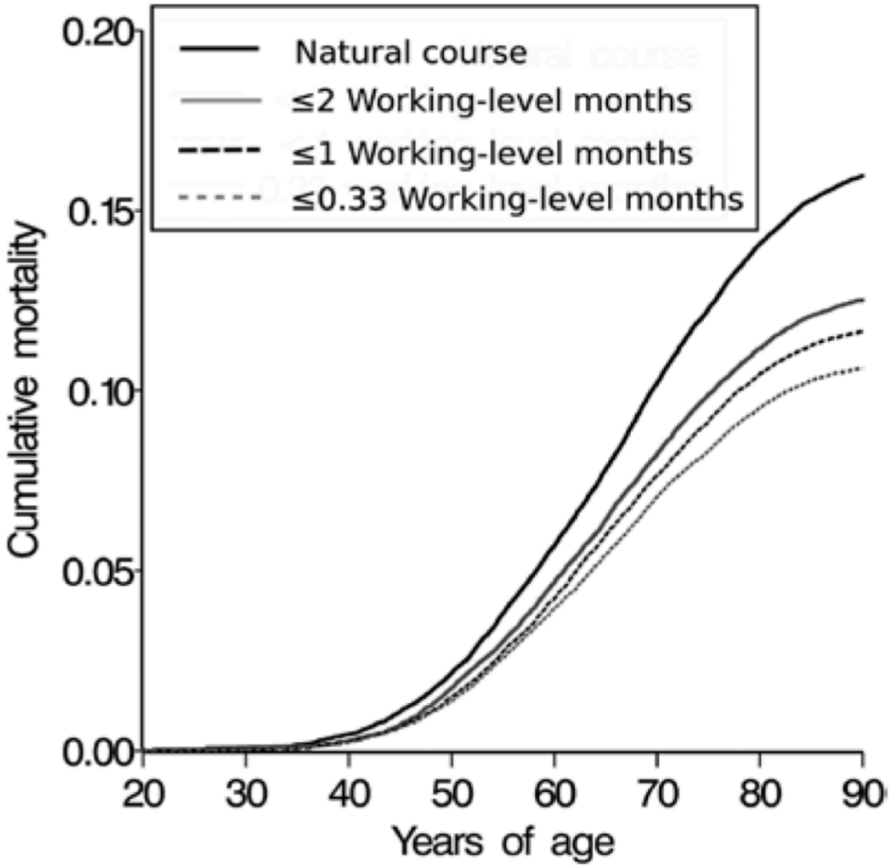


Describing the world under interventions

Tools for causal inference

Potential outcomes framework

Again, focus on risk

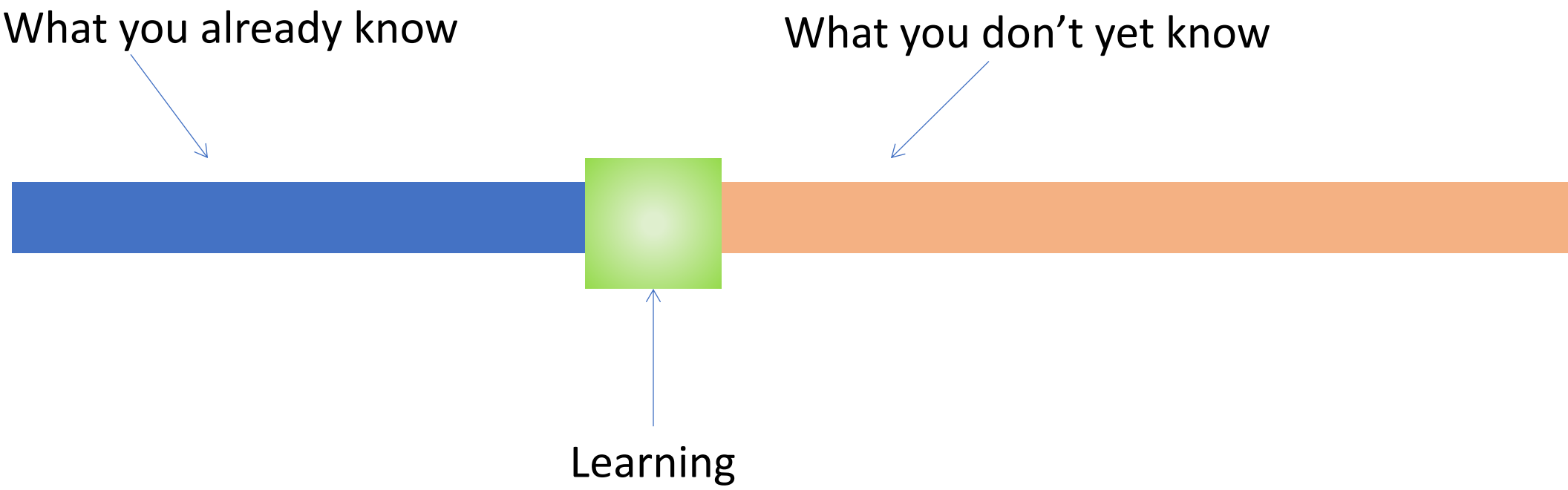


A note on learning methods

Narrow focus on a single parameter (risk) and a specific set of tools.

But learning about risk and this specific set of tools is not the point of the course.

The point is to develop the skills you need to learn *any* new method.


What you already know
What you don’t yet know
Learning

Course Structure

Course web tools

1. **Course website**: course info, schedule, lecture materials
2. **Microsoft Teams**: communication with teaching team, download exercises and data, submit exercises
3. **Sakai**: access recorded lectures

Feedback

We welcome your feedback throughout the course.

By email to the instructors or the TAs.

Anonymously through the form on the website.

Final course evaluation (through UNC evaluation system).

Closing thoughts

Final thoughts

You are responsible for your learning.

Some day very soon, if it hasn't happened already, you will be the "epidemiologist in the room."

What can you learn during this program to prepare you?

Advanced Epidemiologic Methods

EPID 722
Spring 2021
UNC – Chapel Hill
jessedwards@unc.edu