

# Tutorial 4: Spotting and improving biased sampling

ADS2

Semester 1 2023/24

This week's session is a tutorial, rather than a programming practical. You will be assigned to a tutorial group and given a time slot and classroom for a 1-hour tutorial.

Remember that the facilitators are there to guide discussion. As a group, you are responsible for starting, developing and concluding the discussion.

## Learning Objectives

- Explain the concept of sampling bias
- Give examples of sampling biases that can occur
- Design data collection procedures that avoid sampling bias

## Before the class

We will go through several examples of experimental designs, and you will be asked to examine them for possible sampling bias and suggest better ways of designing the study. You can have a look at the discussion questions below, but there is no need for any additional reading apart from this document.

## In class

Discuss the following:

### 1. Zika virus infections in Canadian travellers

#### Background

Zika is a viral infection spread by mosquitoes. There was a recent major outbreak between 2015 and 2017, starting in Brazil and spreading through other parts of Central and South America.

A study by Boggild et al. (CMAJ, 2017) aimed to document cases of Zika in Canada for a year starting in October 2015.

The study sample consisted of people who came to a clinic within the Canadian Travel Medicine Network because of ill health upon returning from a trip to other countries in the Americas. Depending on their symptoms, patients were tested for Zika or other tropical diseases, such as Dengue fever (and sometimes for several diseases).

Some of the main findings of the study were:

- Among travellers returning from the Americas, Zika infection is as common as dengue infection (41 patients in one year for each).
- A Zika infection can lead to complications, e.g. of neurological nature. Those complications were found more commonly in infected Canadians returning from the Americas than previously reported in areas where Zika is common.
- There are two different tests to determine Zika infection. In 83% of cases, patients tested positive in one of the tests but not the other, showing that it is important to use both tests together.

### Questions

- Based on the study design, do you think there may be sampling bias?
- If there is sampling bias, will it likely affect all of the conclusions of the study, or just some of them?
- If there is sampling bias, how could it have been avoided?

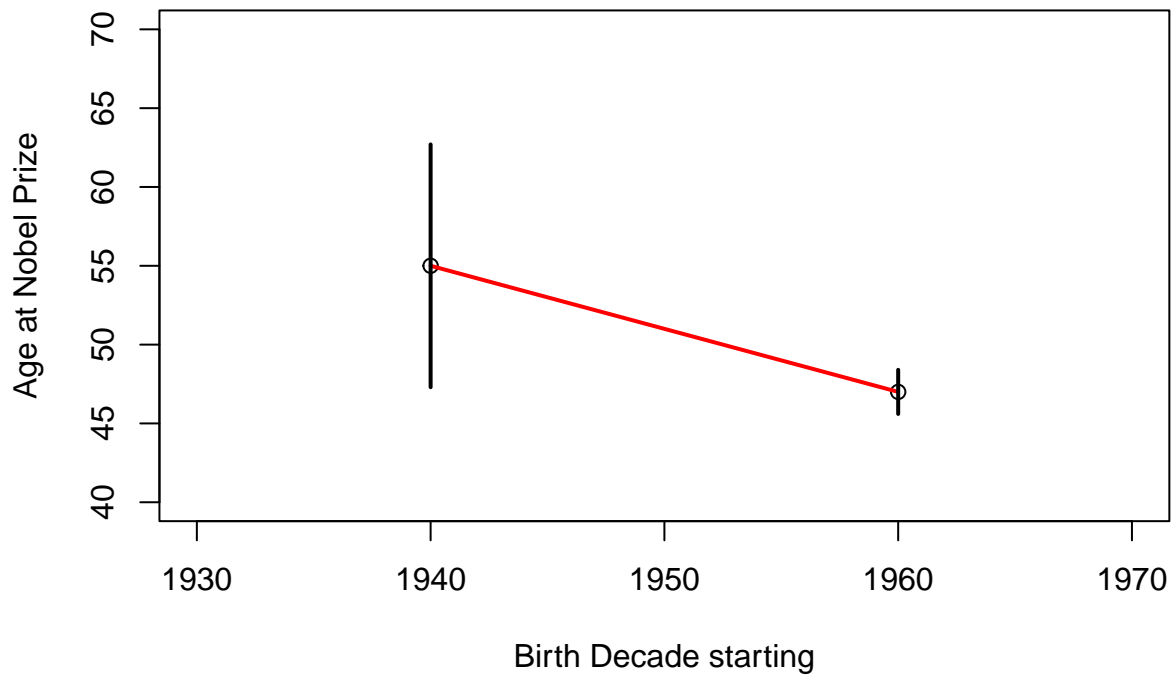
## 2. Age of Nobel Laureates

### Background

Is the age of Nobel Laureates constant over time? A survey in 2012 recorded the age of all Nobel Laureates in Physiology and Medicine at the time they received the Nobel Prize.

Here are the statistics for two groups of laureates:

- Nobel laureates born in the 1940s:
  - Average age: 55
  - Standard deviation: 7.7
- Nobel laureates born in the 1960s:
  - Average age: 47
  - Standard deviation: 1.4



A t-test for difference in age gives a p-value of  $p = 0.0001$ , suggesting that Nobel laureates are getting younger.

### Questions

- Do you agree with this conclusion?
- How was the data sampled? Do you think this is a biased way of sampling and if so, how could it be improved?

### 3. Drinking Whisky to live longer

Grace Jones was Britain's oldest person until she died in June 2019, at the age of 112. According to newspaper reports, her secret was drinking a small amount of whisky every night. Should you do the same?

---

Originally created by MI Stefan in 2019, CC-BY-SA 3.0

Last update by DJ MacGregor in 2023