Group 14: Yukiko Nakada, Victoria Ngo, Kalp Patel, Pam Potempa-Rivers
EPI 569: Concepts and Methods in Infectious Disease Epidemiology
Dr. Lopman, Dr. McQuade, Dr. Nelson
TA: Jessica Rothman
November 17, 2024

# Check-In 1

## Preliminary data check

A concern for the data collected for the Fall Fever outbreak of 2024 is the substantial amount of missing data. We see the most significant amount of missing data entries from case IDs 74 through 91, exposure columns greater than exposure one, and columns related to symptoms where no symptoms were reported. In traditional epidemiology, missing data poses an issue of bias – often towards the null, falsely indicating that the association is less strong than it truly is. While we were not the architects of the designing this study, to mitigate risk of missing data we would have attempted to make each part of the survey required for the participant to submit.

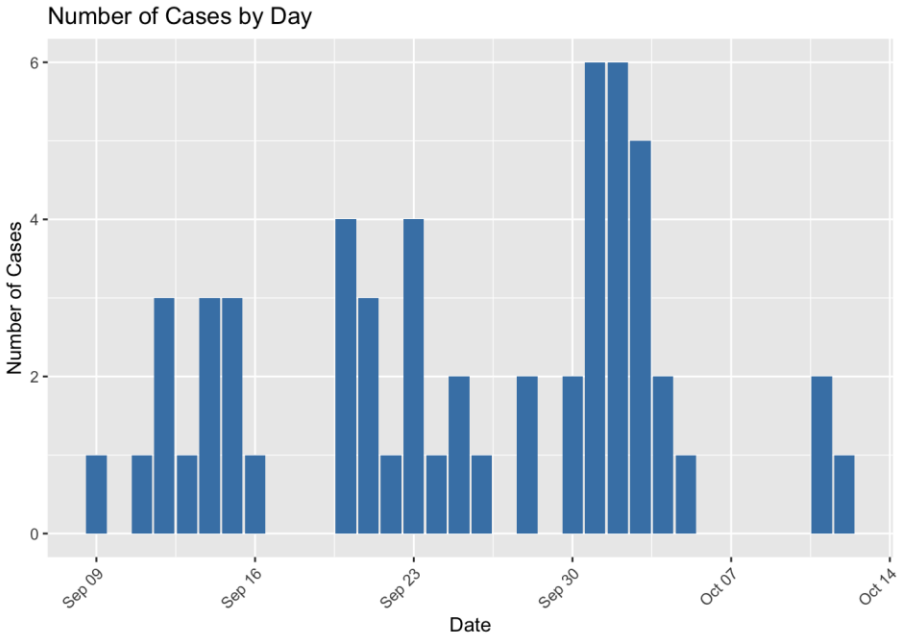
In the analysis phase we will need to clean and restructure data:

1. For our epi curve, we will not include case data that is missing an onset date. We will do this by examining the line list and filtering the cases by onset data as a form of restructuring.
2. To examine the prevalence, we will restructure the line list data by collapsing numerous sections and examining the cases present.
3. We also will join the linelist and transmission pairs files to create one master document.

In the manuscript phase we will:

1. Generate an epidemic curve to visualize the outbreak's progression, interpret key characteristics, and discuss how this informs the natural history and dynamics of Rollins Fall Fever.
2. We will report the prevalence and impact of missing data on analysis and interpretation.
3. Use simulation modeling to evaluate the effects of potential interventions and explore scenarios related to susceptibility and transmissibility using computational approaches.

Case histogram




Potential questions to address

**Question 1:** When was the peak of the outbreak?
**Question 2:** What is the average duration of the infection period? How do we see this in our serial interval? How might this impact our R0?
**Question 3:** What is the basic reproduction number (R0), and how does it change over time ($R_e$)?
**Question 4:** How would the dynamics of an infectious disease outbreak change if individuals no longer acquire complete immunity after infection?
**Question 5:** What proportion of those infected are asymptomatic? What proportion are symptomatic? What does this conceptually indicate about our incubation period, latent period, serial interval, infectious period? What could this mean for containment and control efforts?

Assignment of Roles

**Question 1. Epi Curve** – Pam

**Question 2. Natural History** – Yukiko

**Question 3. Calculating R** – Victoria

**Question 4. Simulation Modeling** – Kalp

**Question 5. Evaluating susceptibility and transmissibility** – Pam