Epidemiology 101

Kimberly J. Johnson, MPH, PhD
Associate Professor, Brown School
Washington University in St. Louis
kijohnson@brownschool.wustl.edu
Open classroom outline

• **Part 1:** What is epidemiology?
• **Part 2:** Measuring disease in populations
• **Part 3:** Understanding causes of disease in populations
• **Part 4:** Controlling disease in populations
• **Part 5:** COVID-19 frequently asked questions
Part 1: What is epidemiology?
Epidemiologists help take care of populations

"I'm sorry, Dr. Etheridge is an epidemiologist, not a general practitioner. But maybe if you get a group together, he would see you."
Public health

• Public health is the science of protecting and improving the health of people and their communities.

• Epidemiology is one of several disciplines in public health including:
  • social and behavioral sciences
  • biostatistics
  • environmental health science
  • health services administration and policy

https://www.cdcfoundation.org/what-public-health
The formal definition of epidemiology

The study of the *distribution* and *determinants* of health-related states or events in *specified populations*, and the application of this study to *control of health problems* (Last 2001, *Dictionary of Epidemiology*)
Epidemiologists gather data to understand:

1. Where disease is occurring (person, place, time) in populations
2. Why it is occurring in populations
   - Modifiable risk factors
   - Non-modifiable risk factors
3. How to prevent and control disease in populations

1. Descriptive Epidemiology (person, place, time)
2. Analytic epidemiology (Why)
3. Control (Interventions)
Core assumptions that guide epidemiology activities

1. Health outcomes are not randomly distributed in a population

2. The causes of health outcomes are multi-factorial
1. Health outcomes are not randomly distributed in a population

Cases 4/3/2020

Graphic courtesy of the New York Times
2. The causes of health outcomes are multi-factorial
Part II: Measuring disease in populations
How do epidemiologists measure disease in populations?

• Surveillance systems
  • Cancer
  • COVID-19 reporting

• Surveys
  • Nationwide
  • Subgroups

• Administrative databases
  • Health claims databases
  • Electronic health records
Cancer Surveillance

• Malignant diagnoses collected from entire US population

• CDCs National Program of Cancer Registries and National Cancer Institute’s SEER program

• Data are collected from a variety of sources to ensure completeness

https://www.cdc.gov/cancer/npcr/about.htm
Cancer Surveillance

• Trends in cancer over time

• Provide data to identify high risk groups for developing and dying from cancer

• Help with planning for cancer control programs

• Helps with deciding on resource allocation for control

• Advances our knowledge about the cancer burden in the U.S. and globally

https://www.cdc.gov/cancer/npcr/about.htm
COVID-19 reporting (US)

- CDC standardized case report form
- Electronic reporting to CDC

How do epidemiologists calculate disease frequency?

- **Prevalence** is the number of *new* and *existing* cases of disease in the population at any given time
  - % of people in a population right now with diabetes

- **Cumulative Incidence** is the number of *new* cases of disease that develop in the population during a defined period
  - The number of new cases of cancer diagnosed per Missouri population in 2020

- **Mortality rate** is the number of deaths in a population during a defined period
  - The number of deaths per Missouri population in 2020

- **Case fatality rate** is the number of deaths among those who have the disease
Case fatality rate

\[
\text{Case fatality} \, (\%) = \frac{\text{No. of deaths from disease}}{\text{No. diagnosed with disease}} \times 100
\]

• Accurate case-fatality rates depend on accurate counts of
  • Death (the numerator)
  • Diagnoses (the denominator)

• Testing ability impacts the accuracy of the denominator and the numerator

• Comparison between geographic locations is challenged by differences in ability to count:
  • Individuals with disease
  • Deaths due to disease
Part III: Understanding causes of disease in populations
Disease causation is multifactorial

- Not everyone who smokes gets lung cancer
- Not everyone who is infected with COVID-19 gets sick
Rothman’s causal pie model

• Envisions disease as a pie where each piece of the pie are causes of disease that must come together for disease to occur in a person

The full pie is a **sufficient cause** for the disease to occur in a person
Each pie piece (i.e., A-P) is a **component cause**
A component cause that is part of every person’s sufficient cause (i.e., A) is a **necessary cause**

Example: COVID-19 example

Mail exposure (hypothetical)

Driver sneezes on package

You touch your eyes, mouth or nose

You touch same spot on package

Virus (on package)

Solution for both scenarios is to NOT acquire all the component causes that make up a sufficient cause. For example, don’t touch your eyes, mouth or nose after touching the mail or subway bar.

Subway exposure (hypothetical)

Person sneezes on subway bar riding before you

You touch your eyes, mouth or nose

Virus (on bar)

Touch metal bar at spot where sneeze landed
Study designs to identify risk and protective factors for disease

1. **Experimental studies** - Researcher assigns intervention
   - Randomized controlled clinical trials
   - Participants are randomized to treatment vs. placebo

2. **Observational studies** - ’nature’ assigns intervention
   - Cohort
   - Case-control
   - Cross-sectional
Cohort study (Vaccination)

Exposed: Vaccinated

Unexposed: Unvaccinated

Measure Infectious disease (ID)

4/14/2020 Disease free

4/14/2021 disease & no disease

 TRACK PARTICPANTS FORWARD IN TIME TO SEE IF THEY GET THE ID

8/10 = 80% with ID

2/10 = 20% with ID
Calculating risk for disease from studies: The relative risk (RR)

\[
RR = \frac{\text{Disease incidence in exposed}}{\text{Disease incidence in unexposed}}
\]

Disease incidence in vaccinated = 2/10 or 20%

Disease incidence in unvaccinated = 8/10 or 80%

\[
RR = \frac{20}{80} = 0.25
\]

Interpretation: The disease incidence in those who were vaccinated was 0.25 times (75% lower) than the disease incidence in those who were not vaccinated.
Considerations for interpreting epidemiologic data

- **Study design**
  - Experimental results from randomized controlled trials are considered the gold standard
  - Anecdotal case reports are useful for generating hypotheses but are not evidence for cause and effect

- **Study size**
  - Results from larger studies are considered more accurate than smaller studies

- **Sampling of participants**
  - Random sampling of the population is more accurate than other methods such as studies that are solely based on who volunteers

- **Confounding**
  - Study result may be due to another factor
    - A lighter in a pocket is associated with an increased risk for lung cancer.
Part IV: Controlling disease in populations
Epidemiological data allows us to make informed decisions about control measures.

- Evidence-based recommendations for:
  - Breast cancer screening
  - Vaccinations for infectious disease
  - Education and policies regarding smoking
  - Social distancing to flatten curve
Mortality rates from lung cancer

Tobacco Use in the US, 1900-2006

*Age-adjusted to 2000 US standard population.

Social distancing for the control of COVID-19 infection

- Effort to reduce person-to-person transmission rate
- Social distancing and stay at home recommendations reduce transmission rates
- Keeping health system from being overwhelmed will help save lives

Part V: COVID-19 FAQ
COVID-19 epi FAQs

We’re starting to hear rumblings of peaks approaching and declines in new case rates, do you see hope in the newest numbers – for StL, NYC, generally?
NYC COVID-19 Counts

STL COVID-19 Counts

Data source: https://www.stlouis-mo.gov/covid-19/data/
COVID-19 epi FAQs

Comparisons have been drawn to influenza (esp 1918-19) and also to HIV. What lessons do you think are apparent? Are the comparisons fair?
Curves for death in St. Louis and Philadelphia

**Philadelphia**
- First cases reported Sept 17, 1918
- City-wide parade Sept 28, 1918
- Social distancing measures implemented Oct 3, 1918

**St. Louis**
- First cases reported 10/5/1918
- Social distancing measures implemented 10/7/1918

stltoday.com
COVID-19 epi FAQs

African American populations appear to be experiencing mortality from COVID at a rate that greatly exceeds that of other groups. What do you think is going on?
Master of Public Health at the Brown School

Transdisciplinary
Evidence-Based
Health Equity Focused
Top #20 Ranked (U.S. News & World Report 2018)

Visit us online: bit.ly/Brown-School-Grad-Programs
Discover our programs:

Master of Public Health
Master of Social Work
Master of Social Policy

Visit us online: bit.ly/Brown-School-Grad-Programs
Any Questions?