# Disease Detectives

Division B/C

Georgia Tech Event Workshop Series 2024-25



01

**RULES SHEET** 

02

**DIFFICULT TOPICS** 

03

**COMMON QUESTIONS** 

04

**TIPS FROM A VETERAN** 

05

**OTHER FREE RESOURCES** 



#### The Rules Sheet

- No significant changes in nearly 6 years...
- 3 main sections
  - Background and Surveillance
    - Free points!
  - Outbreak investigation
    - Use science to solve an outbreak
  - Patterns, controls, and prevention
    - Math and graphs
- Very little about this event is conceptually difficult!
- Challenge is in being able to apply it to nuanced situations



- DESCRIPTION: Participants will use their investigative skills in the scientific study of disease, injury. CALCULATOR: Class II APPROXIMATE TIME: 50 minutes

Each team may bring one 8.5" x 11" sheet of paper, which may be in a sheet protector sealed by tape or laminated, that may contain information on both sides in any form and from any source without any annotations or labels affixed, along with two stand-alone non-programmable, non-graphing calculators

- a. This event addresses three topics related to disease, injury, health, and disability in populations or groups of people. Each part should count approximately equally towards a team's final score. Questions should be process-oriented and involve skills in evaluation and interpretation. Matching pathogens with specific diseases (i.e. What causes X diseases?) or knowledge of signs, symptoms or epidemiologic characteristics such as incubation or latency periods or infectious dose is not part of this event. However, it is appropriate to provide this information as background information and expect competitors to be able
- b. The topics for this event are as follows: i. Background & Surveillance
  - (1) Understand the Clinical Approach (health of individuals) vs Public Health Approach (health of

  - (1) Condestand the Camera Approach (neutron to manytonaday) vs runne resum Approach (neutron populations).

    (2) Understand the instory and development of epidemiology (2) Understand the rotes of epidemiology in public health and the steps in solving health problems (4) Understand the Natural History and Spectrum of Digasaes. Tonderstand in broad cerms the impact of infectious (bacterial, Art, Ingall, protist and prion disease) and noninfectious made the problems.
  - causes of disease (such as accidents, exposures, and toxicities)

    (5) Understand the basic epidemiological and public health terms found in the glossary of CDC's Principles of Epidemiology in Public Health Practice (e.g., outbreak, epidemic, pandemic,
  - Surveillance, risk, vector, etc.)

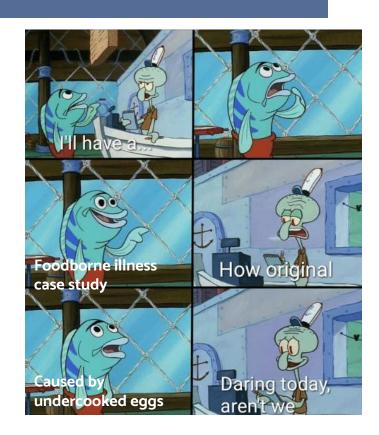
    (6) Understand the role of Surveillance in identifying health problems, the 5-Step Process for Surveillance, and the types of surveillance and the attributes of a surveillance system
  - Outbreak Investigation Analyze actual or hypothetical outbreaks given in case see
  - (1) Analyze actual or hypothetical outbreaks given in case scenarios (2) Understand Experimental and Observational studies and the Types of Epidemiological Studies (e.g., case control, cohort, ecological, cross-sectional). Know the advantages and disadvantages of each. Recognize various fundamental study designs and identify which is appropriate to use in analysis of presented outbreak scenarios
  - (3) Identify the Steps in an Outbreak Investigation and how they guide hypothesis generation
     (4) Identify the problem using person, place, and time triad to formulate case definitions
  - (5) Interpret epi curves, line listings, cluster maps, subdivided tables, PFGE gels, SNP mapping and the PulseNet concent Understand the agent, host, environment triad and chain of transmission
  - (7) Evaluate data by calculating and comparing simple rates and proportions such as attack rate, relative risk, odds-ratio, and explain their meaning. Determine whether presented data support hypotheses of disease within scenarios, and revise hypotheses as appropriate.

    (8) Apply the Bradford Hill Criteria for Verifying the Cause of presented outbreaks, Compare the
  - accuracy of Bradford Hill criteria, Koch's and Evan's postulates, and newer causality models
  - such as Directed Acyclic graphs, Sufficient/component cause models, and GRADE methods

    (9) Understand the concept of herd immunity. Be able to calculate and interpret herd immunity threshold, basic and effective reproductive number

= ©2025-C15





- Case studies are the heart of most Disease
   Detectives tests
- Often you are given a situation then asked to apply principles from all 3 sections to it
  - o For this, you need to use a study
  - But what study to use? And what's the difference between them anyways?

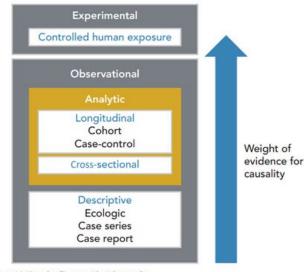


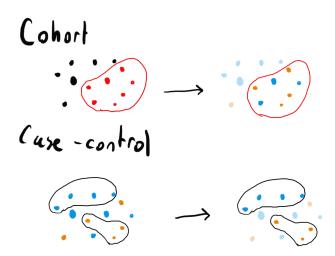
Figure 4-1. Hierarchy of human epidemiologic studies.

#### Cohort

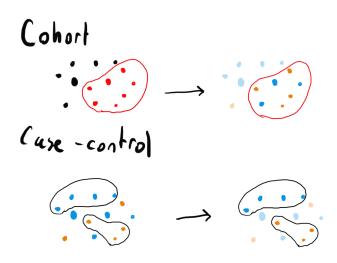
 Sample a group, some will be exposed and some will not, see who develops outcome

#### Case control

- Sample a group with disease and a group without disease
- Increasing controls increases **power** (ability of study to find association if it exists)



- Study types
  - You will have to either choose a study type for the outbreak or identify the study type used
  - Cohort vs Case Control
    - If the population involved is well defined (ie people at a family potluck), it will almost certainly be a cohort study
    - If it is impossible to identify everyone exposed to the agent (ie people who visit the park begin falling ill), only a case control is possible



#### Variations of Cohort

#### Prospective

- Collect participants before outcomes occur
- Attrition bias people may drop out
- Participants must not have outcome at time of data collection, but must be at risk of developing outcome

#### Retrospective

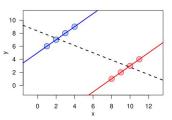
- Collect participants after outcomes occur, use records to determine exposure status
- More risk of bias, but cheaper and easier
- May not have proof of exposure status

#### Variations of Case-control

- Matched / matched-pair
  - Used to control for 1 factor of confounding
  - For each case, choose 1 or more controls with the same condition (ie age)
- Nested case-control
  - Take cases/controls from a cohort
  - Used when looking at everyone in the cohort would be unreasonable ie running expensive genetic tests to see association between gene and cancer
- Case-cohort
  - After a cohort study ends, compare people with outcome to random sample (control group) of people in the cohort
  - Control group can include cases!

#### Biases

- Confounding
  - Something is associated with both the cause and effect you are trying to study
- Sampling bias
  - Selection bias
    - Incidence bias people who have mild/severe cases of disease are excluded
- Berkson's bias / Berksonian bias
- Simpson's paradox



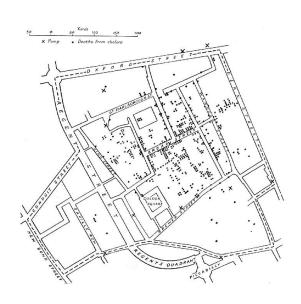
#### Biases

- **Immortal time bias** A type of misclassification assigning a patient to exposed group before they have been exposed
- Recall / response bias recall is when patient remembers incorrectly, response is when patient lies
- Length time Increase time patient is aware of disease without really improving prognosis
- Lead time bias Patients with mild/serious disease not caught by screening

#### Less common study types

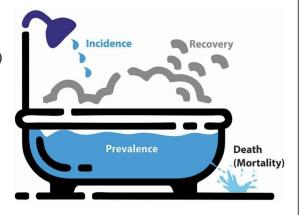
- Cross-sectional study
- Ecological study
  - John Snow's cholera study
  - Ecological fallacy
- Time series
- Case series
- Quasi-experimental
- Experimental

Only experimental studies can find causation! All others can only find correlation.



#### Measures of disease spread/virulence

- Prevalence vs incidence
- Mortality vs morbidity
- Can be proportions (0.0003) or ratios (3 per 10,000 people)



#### Measures of correlation in cohort studies

Risk and relative risk

	Listeria +	Listeria -	
Ate deli meat	5	30	35
Did not eat deli meat	1	45	46
	6	75	81

Risk (ate deli meat) Risk (did not eat deli meat)

5/35=0.143 1/46=0.022

Relative risk: 0.143/0.022=6.5; Those who ate deli meat were 6.5 times more likely to develop listeria

#### Measures of correlation in case-control studies

#### Odds ratio

	Listeria +	Listeria -	
Ate deli meat	5	30	35
Did not eat deli meat	1	45	46
	6	75	81

Odds (ate deli meat) Odds (did not eat deli meat)

5/1=5 30/45=0.667

Odds ratio: 5/.667=**7.496**; Those who developed listeria were **7.5 times** more likely to have eaten deli meat.

Generalizing the calculations...

RR:  $\frac{a/(a+c)}{b/(b+d)}$ 

	Listeria +	Listeria -	
Ate deli meat	а	b	a+b
Did not eat deli meat	С	d	c+d

b+d

a+c

OR: a/c

a+b+c+d

#### Risk ratio (RR) vs Odds ratio (OR)

- OR is an approximation of RR
  - $\circ$  If RR = 1, OR = 1
  - Otherwise OR will be further from 1 than RR
  - If event is rare, OR will be a good approximation of RR (rare disease assumption)
- ...so why use OR for case-control studies?
  - Risk ratio requires "people at risk" as denominator
  - In a case control, you can't know how many people are at risk

- The **basic reproductive number R**<sub>o</sub> represents the number of infections produced by a typical infected individual
  - Estimated to be ~4 for COVID (lots of variation tho)
- The **effective reproductive number R or R<sub>t</sub>** represents the number of infections produced by a typical infected individual *considering susceptibility* 
  - $\circ$  R<sub>t</sub>=R<sub>o</sub>x; x is the proportion of susceptible individuals
  - If half of the population is vaccinated:  $R_t = 4^*.5 = 2$
- Herd immunity threshold is the proportion of individuals that must be vaccinated to stop proliferation of the disease
  - Herd immunity threshold =  $1 1/R_0$



- Confidence intervals show a range of likely values for some number
- Interpretation of 95% CI
  - If the study was done repeatedly, 95% of the reported confidence intervals would include the 'true' value

#### 4: Univariate analysis of patient risk factors for Legionnaires' disease in visitors to the Aquarium\*

Variable	Cases (n = 104)	Controls ( <i>n</i> = 201)	Odds ratio (95% CI) <sup>†</sup>
Seeing a doctor regularly	65	119	1.15 (0.71–1.87)
Seeing a doctor for a problem of lungs, heart or diabetes	38	60	1.35 (0.82–2.23)

#### Sensitivity and specificity

- Tests for a disease can be positive or negative, as well as true or false.
- Sensitivity is the chance a diseased person will get a positive result (TP/TP+FN)
- Specificity is the chance a healthy person will get a negative result (TN/TN+FP)
- Sensitivity and specificity have an implicit tradeoff!
  - Values can be adjusted based on the test (threshold value)

	RDT positive	RDT negative	
PCR positive	True Positive	False Positive	
PCR negative	False Negative	True Negative	

RDT (rapid test) is compared against a highly accurate standard, PCR

Rules explicitly state that knowledge of specific diseases should be not tested for this event!

So why do we care about trivia?



- Tests will almost always expect some level of background knowledge of things not explicitly stated in rules
- Tests may ask for examples
- These are good things to know / have on your notes if you have room (see following slides for more details)
  - Foodborne illnesses
  - History
  - Vectorborne diseases
  - General pathogen knowledge

Studying these should come **after** you already have all the basic knowledge nailed down! You do not need to know any of these topics in detail, just being aware of them is more than often enough.

- Foodborne illness + food safety
  - Know what foods are associated with each (maybe put on cheatsheet)
  - Good ones to have.
    - Salmonella, Norovirus, E. Coli (Shiga-like and non shiga-like), Listeria, Campylobacter, Cryptosporidium, B. Cereus, Shigella, giardia, trichinella
  - On the test: if in doubt, it's salmonella or E.
     Coli

"Oh no! There's a new E. Coli outbreak!"

Disease detectives:

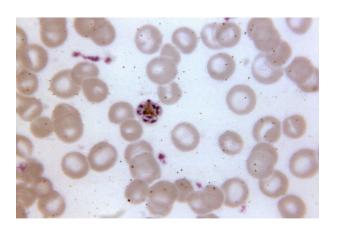


- History
  - Timeline and diseases of major pandemics
    - Plague, HIV/AIDS, SARS and COVID-19, influenza (Do you notice something in common about most of these diseases?)
  - Other historically important diseases
    - TB, Cholera, smallpox (and any disease that we have a vaccine against), malaria
    - Non-infectious diseases: 4 big pollution diseases of Japan
  - Important studies
    - John Snow's Cholera study
    - Doll and Hill's British Physicians study
- Check out this page: <u>History of epidemiology CDC</u>





- Vectorborne diseases
  - Mosquitoes
    - Malaria, dengue, zika virus
  - Ticks
    - Lyme, RMSF
  - Others
    - Chagas disease kissing bug
    - Sleeping sickness tsetse fly
- Why do we care about these?
  - Significant cause of human morbidity
  - Huge impact on communities that lack sufficient healthcare resources
  - Climate change is predicted to make some of these more prevalent



- Pathogens
  - Viruses retroviruses vs general viruses
  - Bacteria
    - Gram +/-/none
    - Antibiotic resistance
  - Parasites
    - Can be macroscopic (dracunculiasis) or microscopic (malaria)
  - Prions
    - UK Beef BSE / TSE outbreak
  - Fungi
    - Rare to see on tests
    - Mostly dangerous to people with weakened immune system



All of the following questions have been pulled from past YJI exams (which can be found on our website) or the Text Exchange on SciOly Wiki

#### Question 1 - Case Study

#### Azure-Sky's test from SSSS 2023

On August 3, 2022, 209 students of Disease Detective Academy headed toward the dining hall. However, a few days later, many students reported symptoms of diarrhea, fever, and stomach cramps. The school has established this as an outbreak.

26. Write a case definition for this outbreak. (4 pts)

After a few hours, you conduct a survey where participants with and without the disease report what foods they ate. For each food, here is the data.

33. What type of study are you conducting, and why? (3 pts)

(Continued on next page)

### Question 1 - Case Study

	Ate (ill)	Didn't eat (ill)	Ate (not ill)	Didn't eat (not ill)
Sweet Mango	82	41	17	69
Steamed Mushrooms	86	37	63	23

Calculate the risk measure for the two foods.

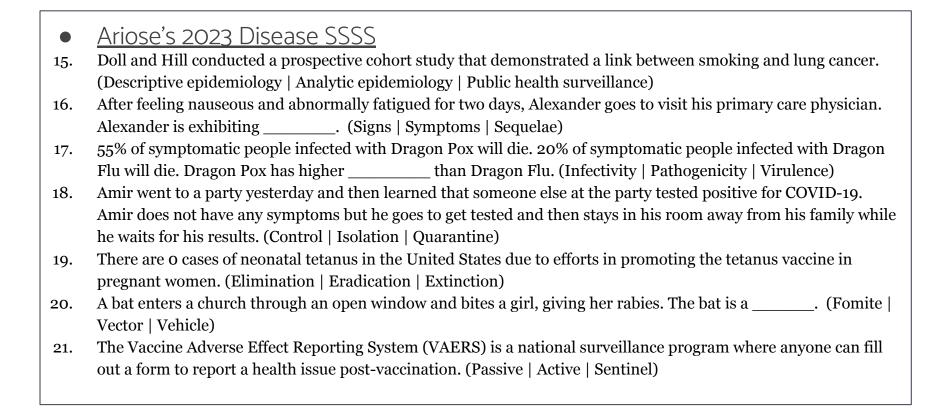
Which food likely caused the outbreak, and how do you know?

Name 3 ways to prevent future foodborne illnesses.

OR: Odds for case/odds for control

(82/41)/(41/69)=8.118

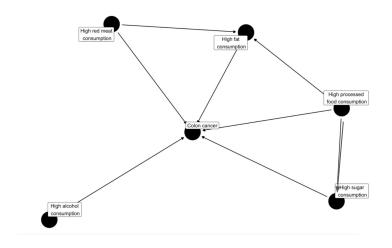
(86/37)/(63/23)=0.849



#### MITSO 2024

The following is a directed acyclic graph (DAG) for the risk factors of colon cancer.

Let's say you only wanted to look at high sugar consumption. Why would it be especially important to control for high processed food consumption?

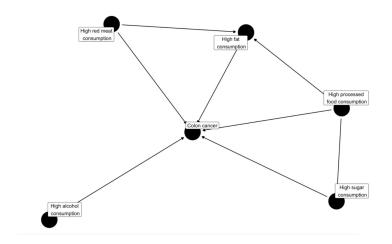


#### MITSO 2024

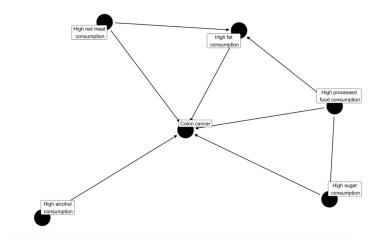
The following is a directed acyclic graph (DAG) for the risk factors of colon cancer.

Let's say you only wanted to look at high sugar consumption. Why would it be especially important to control for high processed food consumption?

Processed foods have high sugar in them

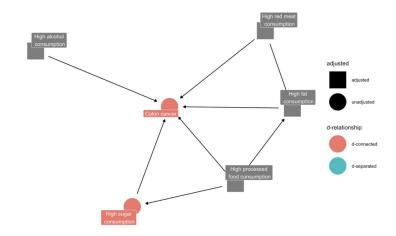


• MITSO 2024 Redraw the DAG from 5 with the minimally sufficient adjustment set if we were only interested in high sugar consumption. Use boxes for controlled factors and circles for factors you are not controlling. You can ignore risk factors for colon cancer that were not part of your original study.



MITSO 2024
 Redraw the DAG from 5 with the minimally sufficient adjustment set if we were only interested in high sugar consumption. Use boxes for controlled factors and circles for factors you are not controlling. You can ignore risk factors for colon cancer that

were not part of your original study.



#### Tips from a Veteran

- Know your basic scioly.org knowledge + vocab (or at least have it on the cheatsheet), do not miss out on the free easy points
- Do practice tests! This is the most effective way to study. Try doing all non-trivia questions without your cheatsheet
- There is not that much foundational knowledge in disease, so most of your cheatsheet should be dedicated to disease trivia, vocab, (and possibly statistics).
  - Common trivia I see people get wrong a lot: Antibiotics should not be used for viral diseases or shigatoxin producing E. Coli
- Google "outbreaks" once in a while to see what the current diseases people are talking about are

#### **Additional Resources**

• <u>History of epidemiology - CDC</u>

CIDRAP

**MMWR** 

CDC TEXTBOOK

## THANKS!

