University of Saskatchewan

Test Analysis

Public Health and Preventative Medicine Residents

University of Saskatchewan

Outline

- Statistical power
- How good is my measure?
- Public health impact measures

UNIVERSITY OF SASKATCHEWAN

Statistical Power

- We want sufficient power to be able detect effects and limit type 1 and 2 errors.
  a) Population health is bad for this because we already start my assuming small effects.
- We don’t want to much power because our study will be super expensive.
- If we are considering equity we need to consider subgroup analyses.

Statistical Power

Table of error types		Reality	
		H_0	H_A
Statistical Decision	H_0	Correct inference (True Negative) (Probability = $1-\alpha$)	Type II error (False Negative) (Probability = β)
	H_A	Type I error (False Positive) (Probability = α)	Correct inference (True Positive) (Probability = $1-\beta$)

Statistical Power

- α = Type 1 error
  a) p value (You pick this)
  b) False positive
- β = Type 2 error
  a) False negative
- $1 - \beta$ = power
  a) You can design a study to get close to this value

- Typical study you use α = *0.05 (5%) and 1-*β = 0.80 (80%).
- You want a study that has a 5% chance of committing a false positive and a 20% chance of committing a false negative.

Statistical power, law, medicine

- Law
  a) We can think of the Null Hypothesis and a true negative as similar to the phrase « Innocent until proven guilty »
  b) We want a study that has a very small chance of saying there is a difference when there isn't one (false positive)
  c) We would rather commit a false negative (let someone guilty go) than a false positive (imprison an innocent person)
- Medicine
  a) **Your examples here.**

Statistical Power

- Sensitivity
  - % of people with disease who have a positive test.
  - TP/(TP+FN)
- Specificity
  a) % of people without disease who have a negative test.
  b) TN/(TN+FP)

Table of error types		Reality	
		H_0	H_A
Statistical Decision	H_0	Correct inference (True Negative) (Probability = $1-\alpha$)	Type II error (False Negative) (Probability = β)
	H_A	Type I error (False Positive) (Probability = α)	Correct inference (True Positive) (Probability = $1-\beta$)

How Good is my measure?

- We often want to know how good a measure is at determining a true positive or true negative. Often we do this because we want cheaper or faster test to be available.
- To determine this we need to have a gold standard or at least criterion measure to compare against.
- Examples?

How Good is my measure?

- Once we have a gold standard we can use measure people with and without disease using the criterion (gold standard) and the comparison measure.
- We can then compare the measures to determine
  a) Sensitivity
  b) Specificity
  c) Positive Predictive Value
  d) Negative Predictive Value

University of Saskatchewan

How Good is my measure?

- Sensitivity
  - % of people with disease who have a positive test.
  - TP/(TP+FN)
- Specificity
  a) % of people without disease who have a negative test.
  b) TN/(TN+FP)

- Positive Predictive Value
  - Out of all of the positive findings, how many are true positives
  - TP/(TP+FP)
- Negative Predictive Value
  a) Out of all of the negative findings, how many are true negatives
  b) TN/(TN+FN)

How Good is my measure?

- Sensitivity and specificity tend to be inversely related
  a) As sensitivity increases, specificity *tends to decrease*, and vice versa. Highly sensitive tests will lead to positive findings for patients with a disease, whereas highly specific tests will show patients without a finding having no disease.
  b) Sensitivity and specificity should always merit consideration together to provide a holistic picture of a diagnostic test

How Good is my measure?

- Sensitivity and specificity tend to be inversely related
  a) As sensitivity increases, specificity *tends to decrease*, and vice versa. Highly sensitive tests will lead to positive findings for patients with a disease, whereas highly specific tests will show patients without a finding having no disease.
  b) Sensitivity and specificity should always merit consideration together to provide a holistic picture of a diagnostic test

How Good is my measure?

- Disease prevalence in a population affects PPV and NPV:
  a) When a disease is highly prevalent, the test is better at ‘ruling in' the disease and worse at ‘ruling it out.’
     - Why?
  b) Disease prevalence should be considered when examining diagnostic test metrics.

Public Health Impact Measures

- Number needed to treat
- Attributable risk
- Population attributable risk (fraction)

Number needed to treat

- The number needed to treat (NNT) or number needed to treat for an additional beneficial outcome (NNTB) is an epidemiologic measure used in communicating the effectiveness of a health-care intervention, typically a treatment with medication.
- The NNT is the average number of patients who need to be treated to prevent one additional bad outcome (e.g. the number of patients that need to be treated for one of them to benefit compared with a control in a clinical trial).
- It is defined as the inverse of the absolute risk reduction

https://en.wikipedia.org/wiki/Number_needed_to_treat

Attributable Risk

- **Attributable risk**
  a) Difference in incidence rates between people exposed to some risk factor vs. people not exposed to the risk factor.
- **Attributable risk percentage**
  a) Percentage of an incidence rate that can be attributed to some risk factor.
- **Population attributable risk percentage**
  a) Percentage of an incidence rate in the overall population that can be attributed to some risk factor.

Attributable Risk

- **Attributable risk**
  a) AR = (A/(A+B)) – (C/(C+D))
- **Attributable risk percentage**
  a) AR % = AR / (A/(A+B)) * 100

	Disease	No Disease
Exposed	A	B
Unexposed	C	D

- **Population attributable risk percentage**
  a) PAR % = [((A+C) / N) – (C/(C+D))] / ((A+C) / N) * 100

Attributable Risk

- **Attributable risk**
  a) AR = (A/(A+B)) – (C/(C+D))
  b) AR = (25/(25+140)) – (52/(52+683))
  c) AR = **.08077**
- **Attributable risk percentage**
  a) AR % = AR / (A/(A+B)) * 100
  b) AR % = .08077 / (25/(25+140)) * 100
  c) AR % = **53.31%**

	Heart Disease	No Heart Disease
Smoker	25	140
Non Smoker	52	683

Attributable Risk Percent

- **Attributable risk percentage**
  a) AR % = AR / (A/(A+B)) * 100
  b) AR % = .08077 / (25/(25+140)) * 100
  c) AR % = **53.31%**

	Heart Disease	No Heart Disease
Smoker	25	140
Non Smoker	52	683

This means 53.31% of incidence of cardiovascular disease among smokers is attributable to their smoking.

Population Attributable Risk %

- **Population Attributable Risk %**
  - PAR % = [((A+C)/N) – (C/(C+D))] / ((A+C) / N) * 100
  - PAR % = [((25+52)/900) – (52/(52+683))] / ((25+52) / 900) * 100
  - PAR % = **17.31%**

- This means 17.31% of incidence of cardiovascular disease in the population is attributable to smoking.

	Heart Disease	No Heart Disease
Smoker	25	140
Non Smoker	52	683

Considering these metrics

- Population attributable risk
  a) PAR assumes that the association is causal. I think this had result in it being underused because people are afraid of causality.
  b) There are new fancier PAR methods that allow more flexibility with causal and other assumptions like monotonicity and no bias (https://www.bmj.com/content/360/bmj.k757).
- These measures convey something different than risk and tend to be more understood by the public. They don't necessarily speak to individuals.

University of Saskatchewan

Work