Measures of Association

Mabel Carabali

EBOH, McGill University

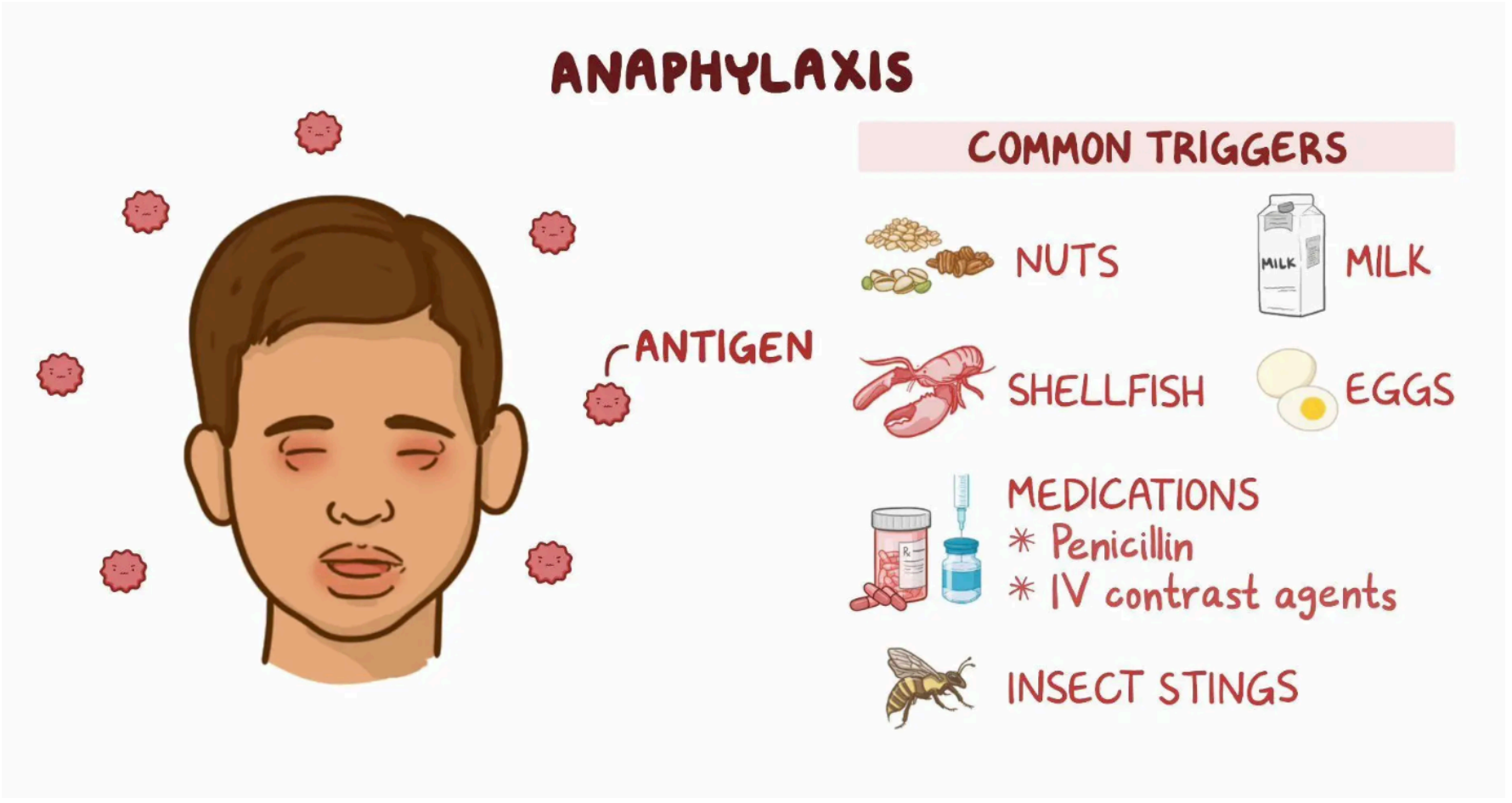
Updated: 2025-09-08

Expected Competencies

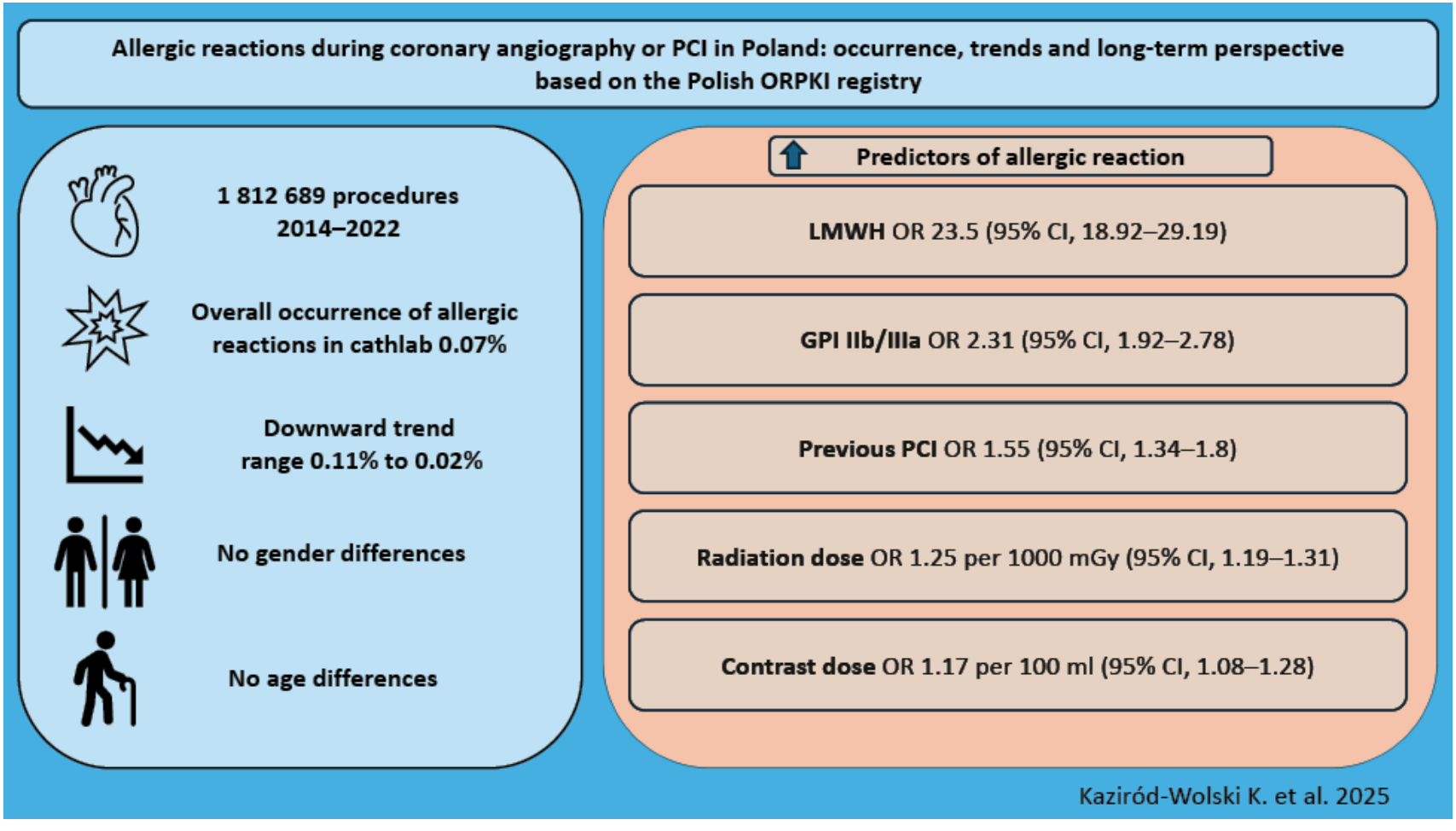
- Knows what are measures of frequency, association and effect.
- Know the difference between relative and absolute measures.
- Recognize and correctly interpret measures of association.

Objetives

- To clarify potential misconceptions about measures of association.
- To introduce the counterfactual and potential outcomes framework.
- To identify advantages and adequate use of different measures.



Anaphylaxis: Clinical sciences. Osmosis from Elsevier



Allergic reactions during coronary angiography or PCI in Poland: Occurrence, trends, and long-term perspective based on the Polish ORPKI registry. Polish Heart Journal (Kardiologia Polska) Vol 83, No 3 (2025)

What's an Association?

Our common objective of epidemiologic research:

- The effect of exposure X on the occurrence of outcome Y
    - But we can rarely observe or even estimate this effect directly.

– It involves the same people at the same time in contrasting exposures, which is impossible.

– We **observe an association** between the *Exposure* and *Outcome* among study subjects, which estimates a population association.

"The observed association will be a poor substitute for the desired effect, if it is a poor estimate of the population association, or if the population association is not itself close to the effect of interest." ME4 (2020)

What's an *Effect* ?

Effect here means the end point of a causal mechanism, i.e., identifying the type of outcome that a cause produces.

EXAMPLE: *"Liver cirrhosis is an **effect** of chronic excessive alcohol consumption".*

- This use of the term effect **merely** identifies liver cirrhosis as ***one consequence of chronic excessive alcohol consumption***.
  - **Compared** to something else (e.g., Abstinence or another level of consumption).
  - Cirrhosis may **Not be the only** effect of of chronic excessive alcohol consumption.
  - May **change** across populations and or over time.

"An effect of some factor is thus relative to the outcomes, to the population, and to the time frame." [ME4 (2020)]

Exposure vs Cause

An exposure (usually denoted as X) is a *potential* causal characteristic, *"a factor that produces an outcome*"

- Could be the *sole* or *compounded* cause [1] of an Outcome. Can be a behavior, a treatment/intervention, a social condition, a health condition, a genetic trait...

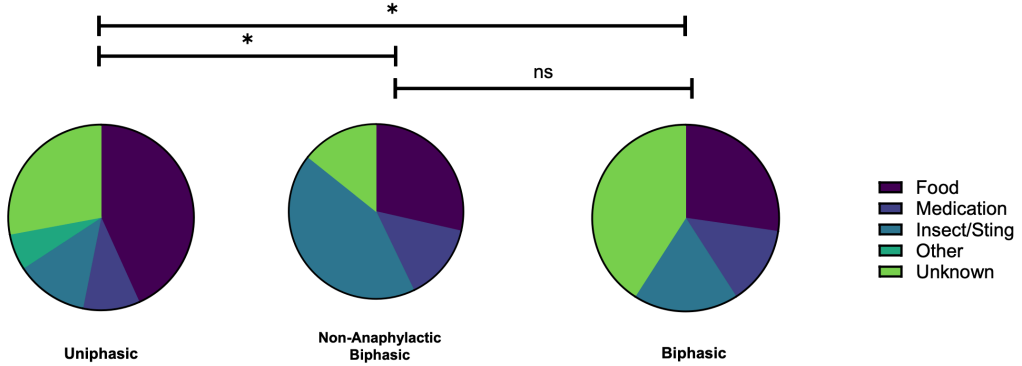
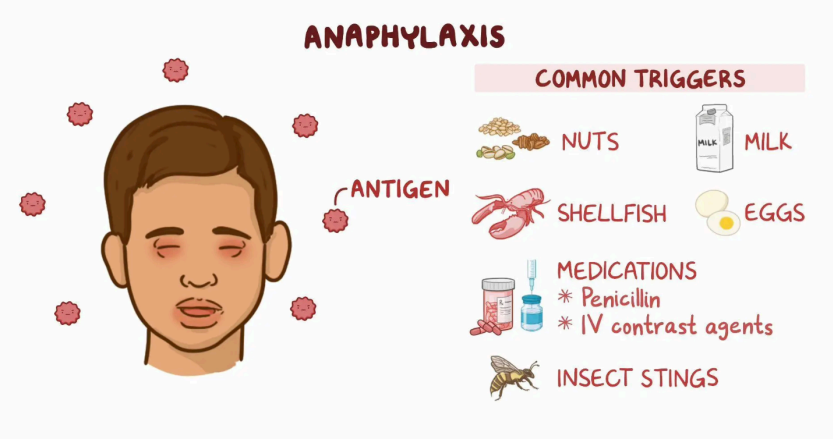


Fig. 1 Reaction profiles. The proportion of anaphylactic triggers across the responder groups is represented by pie charts. The reaction profiles significantly differed between uniphasic responders and non-anaphylactic biphasic and biphasic responders ($p = 0.0179$ and $p = 0.0375$, 2-way ANOVA with Tukey's Multiple Comparisons test). Reaction profiles were comparable between non-anaphylactic biphasic and biphasic responders ($p = 0.8665$, 2-way ANOVA with Tukey's Multiple Comparisons test)

[1] For more on Sufficient Component Causal Framework and Bradford Hill criteria, in this course we focus on potential outcomes and causal DAGs for approaches to causal inference.

- Ellis A.K, et al. (2025) Biphasic anaphylaxis in a Canadian tertiary care centre: an evaluation of incidence and risk factors from electronic health records and telephone interviews

Exposures vs Causes



Anaphylaxis: Clinical sciences. Osmosis from Elsevier

Anaphylaxis Etiology

> *"Common inciting sources may include **exposure** to certain medications, foods, or insect stings."*
>
> *"Occasionally, the offending agent is not identified; these reactions are idiopathic anaphylaxis."*
>
> *"The most common **causes** include bee stings, fire ant bites, latex, and foods (peanuts, tree nuts, fish, shellfish, milk, eggs, wheat, spelt, rye, barley, soy, red meat, and sesame)."*

Anaphylaxis StatPearls NLM

Exposures vs Causes

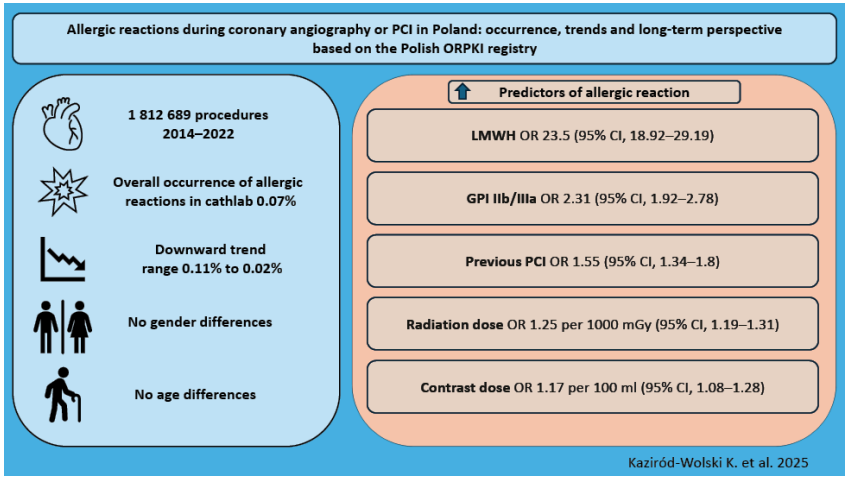


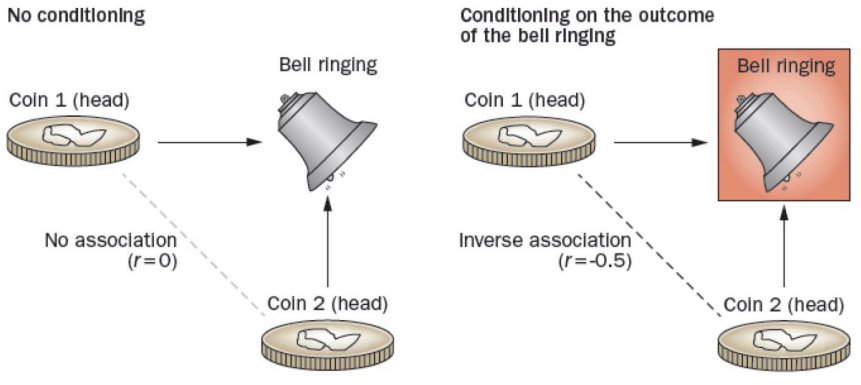
Table 3. Clinical and periprocedural factors affecting incidence allergic reaction

Variables	Univariable	
	OR (95% CI)	*P*-value
Male sex	0.93 (0.83–1.04)	0.21
Age, years	1 (0.99–1)	0.24
Weight, kg	1 (1–1)	0.49
Diabetes	0.87 (0.75–1)	0.06
Previous stroke	1.14 (0.83–1.57)	0.42
Previous myocardial infarction	1.27 (1.12–1.44)	<0.001
Previous PCI	1.37 (1.22–1.55)	<0.001
Previous CABG	0.89 (0.68–1.16)	0.37
Smoking status (active)	0.8 (0.68–0.94)	0.01
Arterial hypertension	0.96 (0.85–1.09)	0.56
Kidney disease	0.68 (0.5–0.92)	0.01
COPD	0.46 (0.27–0.78)	0.003
ACS (yes/no)	1.29 (1.15–1.45)	<0.001
UFH during angiogram	0.67 (0.52–0.87)	0.003
LMWH during angiogram	0.53 (0.2–1.41)	0.21
Cardiac arrest at baseline	1.24 (0.62–2.48)	0.55
Direct transport to cathlab	0.64 (0.41–0.99)	0.045
UFH	0.34 (0.3–0.39)	<0.001
LMWH	49.18 (43.85–55.15)	<0.001

Allergic reactions during coronary angiography or PCI in Poland: Occurrence, trends, and long-term perspective based

Not every Association between Exposure an Outcome is *"Causal"*

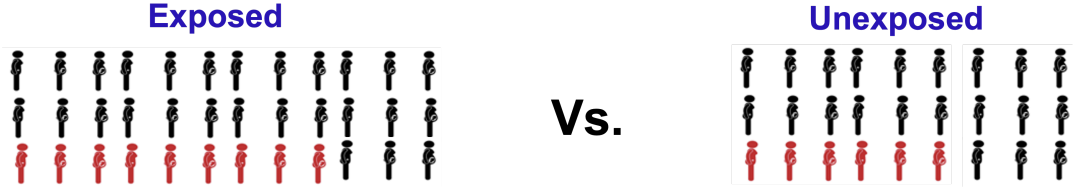
Recall:



At the population level, we assess the effects with measures of occurrence and we **estimate the associations by contrasting such measures of occurrence in the population.**

Absolute vs. Relative measures

- Absolute effect measures are **differences** in occurrence measures.
- Relative effect measures are **ratios** of occurrence measures.



Null Value

Absence of contrast on either the Absolute (Difference) and Relative (Ratio) Scales

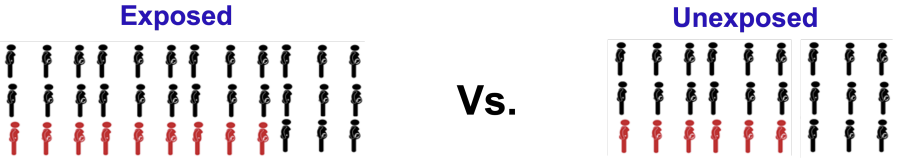
$$R_{exp} - R_{Non-exp} = 0$$

and

$$\left(\frac{R_{exp}}{R_{Non-exp}} \right) = 1$$

Absolute vs. Relative measures

- Absolute effect measures are **differences** in occurrence measures.
- Relative effect measures are **ratios** of occurrence measures.



Sample	Outcome +	Outcome -	Risk Among Exposed	Risk Among Non-Exposed	Risk Difference	Risk Ratio
63	9/36	6/27	0.25	0.22	0.03	1.12

Key elements: Exposure, Outcome & Measures of Occurrence !

The 2x2 Table

- A summary table of observations.

	Outcome		
	Outcome	No.Outcome	Total
Exposed	A	B	A+B
No Exposed	C	D	C+D
Total	A+C	B+D	A+B+C+D

The 2x2 Table

From the previous example:

	Outcome		
	Dead	Alive	Total
Exposed	9	27	36
No Exposed	9	18	27
Total	18	45	63

The measures are:

Sample	Outcome +	Outcome -	Risk Among Exposed	Risk Among Non-Exposed	Risk Difference	Risk Ratio
63	9/36	6/27	0.25	0.22	0.03	1.12

Key elements: Exposure, Outcome & Measures of Occurrence !

Absolute measures: Risk Differences

The RD provides the absolute change in risk

- Indicates how much of the effect is attributable to exposure
- It does not provide information about the magnitude of the shift on the estimates 1 to 16% or 81 to 96%?
- Clinical vs statistical importance
- Public Health Relevance

Relative Measures: Risk Ratios

- Relative measures are popular and practical
- Easier to obtain
  - Dichotomous outcomes!
- Useful in both causal inference and prediction

Interpretability?

- RR's magnitude change according to the coding scheme

Example: (0 to 1) $\neq$ (2 to 1)

Why not both?

- It's possible to see a reduction in absolute estimates, but an increase in relative measures (and vice versa)

These are complimentary estimators! Both tell you something different about the data

- In fact, STROBE and CONSORT guidelines now advise researchers to publish both measures

695 reactions vs 482 = 29% increase vs 20 more cases per 1000 people?

In an epidemiological utopia, researchers would run the model of their choice, obtain relative and absolute estimates, and publish these along with the baseline/background risk

But we live in the real world, so we're more likely to encounter...

...Odds ratios

$$Odds = \left(\frac{P}{1-P}\right)$$

The Odds ratio is the relative contrasts of the Odds among Exposed and the Odds among Non-Exposed

$$OddsRatio = \left(\frac{\left(\frac{P_{exp}}{1-P_{exp}}\right)}{\left(\frac{P_{No-exp}}{1-P_{No-exp}}\right)}\right)$$

...Odds ratios

We are already aware of some key problems with odds and therefore odds ratios

- They overestimate risks
  - While probabilities are bounded [0, 1], **odds** can range from 0 to ∞
- They're not intuitive (except as an approximation of the risk ratio [1])
- And most of the time we care about probabilities, not odds

[1] **When probability is small (<0.10) or given the study design (e.g., case-cohorts) with rare outcomes**

Risk Differences and Risk Ratios

Sample	Risk Among Exposed	Risk Among Non-Exposed	Risk Difference	Risk Ratio	Odds Ratio
63	0.25	0.22	0.03	1.12	1.17
63	0.17	0.15	0.02	1.12	1.15
630	0.017	0.015	0.002	1.12	1.15
630	0.25	0.22	0.03	1.12	1.17

Null Value

Absence of contrast on either the Absolute (Difference) and Relative (Ratio) Scales

$R_{exp} - R_{Non-exp} = 0$; or $\left(\frac{R_{exp}}{R_{Non-exp}}\right) = 1$

Example:

Sample	Outcome +	Outcome -	Risk Exposed	Risk Non-Exposed	Risk Difference	Risk Ratio	Odds Ratio*
63	4/36	3/27	0.11	0.11	0	1	1

- Null Value (Absolute) = (4/36) - (3/27) = 0
- Null Value (Relative) = (4/36) / (3/27) = 1

Note

1. When the RR is above or below the null (>1 or <1) the ORs is FURTHER away form the Null.
2. When the absolute risk in each exposure groups are high, the OR will considerable overestimate the RR.

Some simulated Examples

- Generating 10 data-sets with the *same* structure.

```
set.seed(7042025)
z <- rnorm(500)
e <- matrix(NA,nrow=500,ncol=10) # create an empty matrix to put stuff in
for (i in 1:10) {              # loop 10 times
    e[,i] <- ifelse((rnorm(500))>0.8,1,0) # create a new vector from a binomial
}
e <- as.data.frame(e) # change it into a data frame
names(e) <- c("A","B","C","D","E","F","G","H","I","J") # change the column names

# create a Y matrix with the specification from the questions
y <- ifelse(e==1,rbinom(5000,1,0.65),rbinom(5000,1,0.2))
y <- as.data.frame(y) # change the matrix to a dataframe
# change the names, paste0 says paste these two things together without a space between them
names(y) <- paste0("Y",letters[1:10])
```

Some simulated Examples

Verification of "mean" values across datasets.

```
##        A                B                C                D                E
##  Min.   :0.000   Min.   :0.000   Min.   :0.00   Min.   :0.000   Min.   :0.00
##  1st Qu.:0.000   1st Qu.:0.000   1st Qu.:0.00   1st Qu.:0.000   1st Qu.:0.00
##  Median :0.000   Median :0.000   Median :0.00   Median :0.000   Median :0.00
##  Mean   :0.194   Mean   :0.206   Mean   :0.21   Mean   :0.188   Mean   :0.23
##  3rd Qu.:0.000   3rd Qu.:0.000   3rd Qu.:0.00   3rd Qu.:0.000   3rd Qu.:0.00
##  Max.   :1.000   Max.   :1.000   Max.   :1.00   Max.   :1.000   Max.   :1.00
##        F                G                H                I                J
##  Min.   :0.000   Min.   :0.0   Min.   :0.000   Min.   :0.000   Min.   :0.000
##  1st Qu.:0.000   1st Qu.:0.0   1st Qu.:0.000   1st Qu.:0.000   1st Qu.:0.000
##  Median :0.000   Median :0.0   Median :0.000   Median :0.000   Median :0.000
##  Mean   :0.232   Mean   :0.2   Mean   :0.164   Mean   :0.216   Mean   :0.218
##  3rd Qu.:0.000   3rd Qu.:0.0   3rd Qu.:0.000   3rd Qu.:0.000   3rd Qu.:0.000
##  Max.   :1.000   Max.   :1.0   Max.   :1.000   Max.   :1.000   Max.   :1.000
```

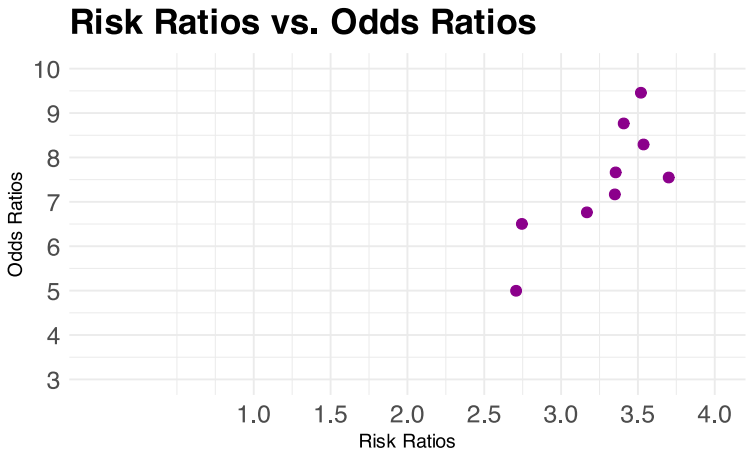
Some simulated Examples

Using "hand calculations" formulas.

```
# Using lapply
tabs <- lapply(1:10,FUN=function(x) table(e[,x],y[,x]))
# With a loop create an empty data frame with the right dimensions
ests <- as.data.frame(matrix(NA,nrow=10,ncol=3))
# Name columns
names(ests) <- c("RD","RR","OR")

# Loop it up
for (i in 1:10) {
x <- tabs[[i]]
ests[i,1] <- x[2,2]/sum(x[2,])- x[1,2]/sum(x[1,]) #RD
ests[i,2] <- (x[2,2]/sum(x[2,]))/ (x[1,2]/sum(x[1,])) #RR
ests[i,3] <- (x[2,2]/sum(x[2,1]))/ (x[1,2]/sum(x[1,1])) #OR
}
```

Some simulated Examples

RD	RR	OR
0.43	3.70	7.55
0.36	2.71	5.00
0.43	3.35	7.17
0.50	3.52	9.46
0.47	3.54	8.29
0.45	3.36	7.66
0.49	3.41	8.77
0.43	2.74	6.50
0.52	3.73	10.50
0.43	3.17	6.76



Some simulated Examples

Using regressions [1] to obtain the estimates (*Rare-ish* outcomes)

```
set.seed(7042025)
yea.dat <- function(n) {
          E <- rbinom(n,1,0.55) #parameters for E
          Y <- rbinom(n,1,0.12) #parameters for Y

    return(data.frame(E=E,Y=Y)) #ask to return a data set with those parameters
}
sim100 <- lapply(1:100,FUN=function(x) yea.dat(400))
summary((sim100[[13]]))  #; summary((sim100[[93]]))
```

```
##        E                Y
##  Min.   :0.0000   Min.   :0.00
##  1st Qu.:0.0000   1st Qu.:0.00
##  Median :1.0000   Median :0.00
##  Mean   :0.5025   Mean   :0.11
##  3rd Qu.:1.0000   3rd Qu.:0.00
##  Max.   :1.0000   Max.   :1.00
```

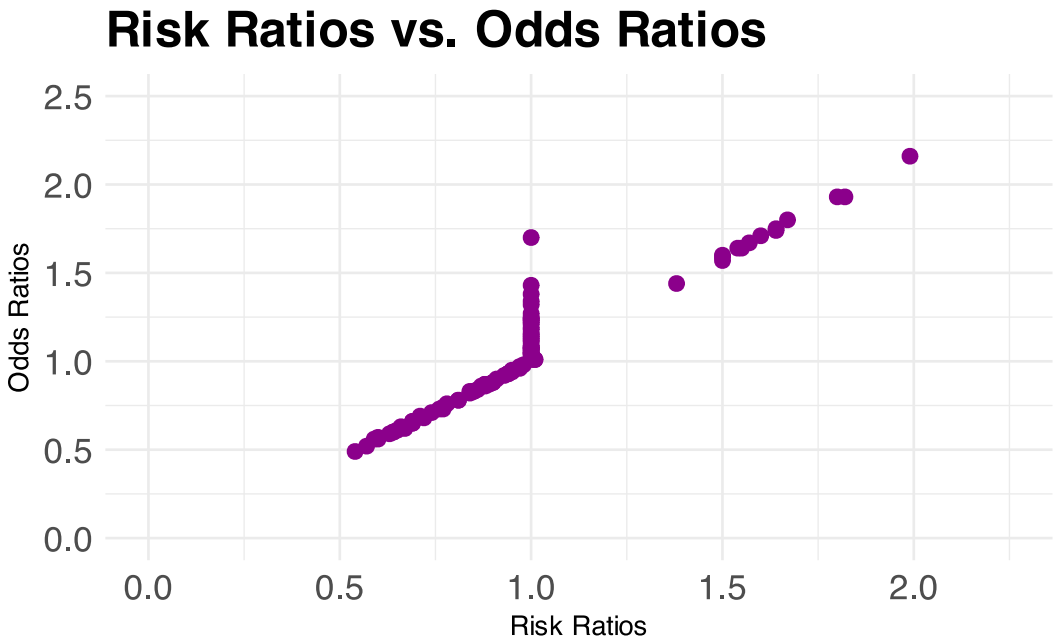
[1] *Some of you may have advanced knowledge on regression analysis but since we have not explained it during the course this resource is only for illustration purposes.*

Some simulated Examples

```
RRs <- sapply(sim100,FUN=function(x) {
   results <- logbin(Y ~ E , data=x)$coef
   return(round(exp(results[names(results)=="E"]),2))
 })

ORs <- sapply(sim100,FUN=function(x) {
   results <- glm(Y ~ E , family="binomial", data=x)$coef
   return(round(exp(results[names(results)=="E"]),2))
 })

sim_RRs<-round(quantile(RRs,probs = c(0.05,0.5,0.95)),2)
sim_ORs<-round(quantile(ORs,probs = c(0.05,0.5,0.95)),2)
```

Some simulated Examples

```
##             5%  50%  95%
## sim_RRs 0.63 1.00 1.80
## sim_ORs 0.59 1.02 1.93
```



Associations Vs Causes

Pearl (2000) uses $Pr(Y = y|SET[X = x])$ to define the probability of an event if the condition $X = x$ were **enforced uniformly** over a population.

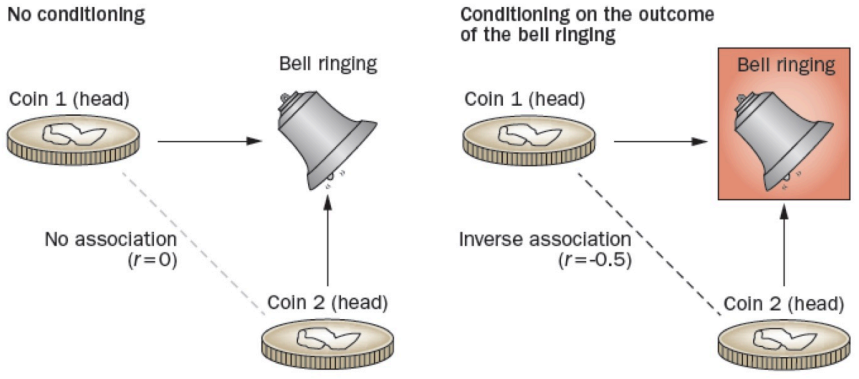
– The key to this definition: **it involves intervention, not observation.**

– Measures of effect can be built based on **SET** notation by creating contrasts of probabilities (or risks) across different X values.

– The What If? Book, express the same notion with $Y^{X=x}$.

Associations Vs Causes

Recall:



Be aware of the difference!

Measures of causal effect

Measures of **causal effect** require a contrast of two **counterfactual** quantities:

- $Pr(Y_i = y_i|SET[X_i = 1]) - Pr(Y_i = y_i|SET[X_i = 0])$

Measures of **association** involve a contrast of two **observed** quantities:

- $Pr(Y_i = y_i|X_i = 1) - Pr(Y_i = y_i|X_i = 0)$

Who are we interested in?

Target Population: The group of people about which the scientific or public health question is asked, in the **relevant etiologic time period.**

“Target Population” in Encyclopedia of Biostatistics, 2005 [Sander Greenland]

> *"The concept of a target population is an informal one, sometimes defined as “the population about which information is wanted” [1] or the “totality of elements which are under discussion and about which information is desired” [4] ...The word “target” emphasizes, however, that this population is not necessarily the same as the one that we end up sampling. The latter population is sometimes called the sampled population [1, 4] or (in epidemiology) the source population [6]". [1] Cochran, W.G. (1977). Sampling Techniques, 3rd Ed. Wiley, New York. [4] Mood, A.M., et al. (1974). Introduction to the Theory of Statistics. McGraw-Hill, New York. [6] Rothman, K.J. & Greenland, S. (1997). Modern Epidemiology, 2nd Ed. Lippincott, Philadelphia.*

Potential Outcomes Framework

We are interested in the effect of exposure $(A = 1)$ on the occurrence of disease $(Y = 1)$

- Suppose everyone in the **target population** of inference is unexposed $(A = 0)$ and we can observe the distribution of Y in the population.

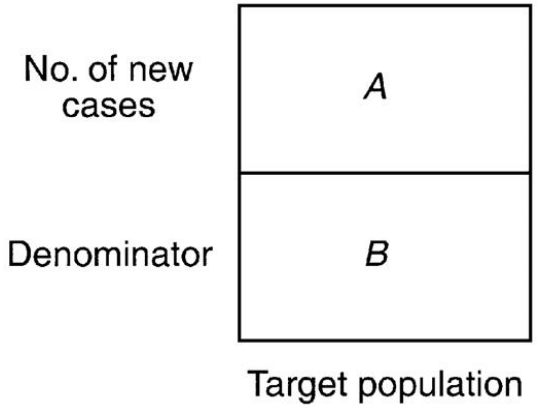
- We would like to also observe the distribution of Y had these same people been all exposed $(A = 1)$

    - This is "counter-to-fact", and we call this condition the counterfactual
    - Each individual has their own **counterfactual** exposure
    - (what would have happened to me if ...)

Potential Outcomes Framework

- We can never observe both conditions in the same population (or individual).
    - That is, we cannot observe the distribution of disease under $A = 1$ and $A = 0$ within the same time period in the same cohort.
    - Thus we need to make an estimate under the condition we do not observe.
- To do so, we use a **substitute population.**
- Our goal is to choose a substitute population that will best mimic **what would have happened to the target population had they experienced the other exposure condition.**

Estimating Causal Effects [1]



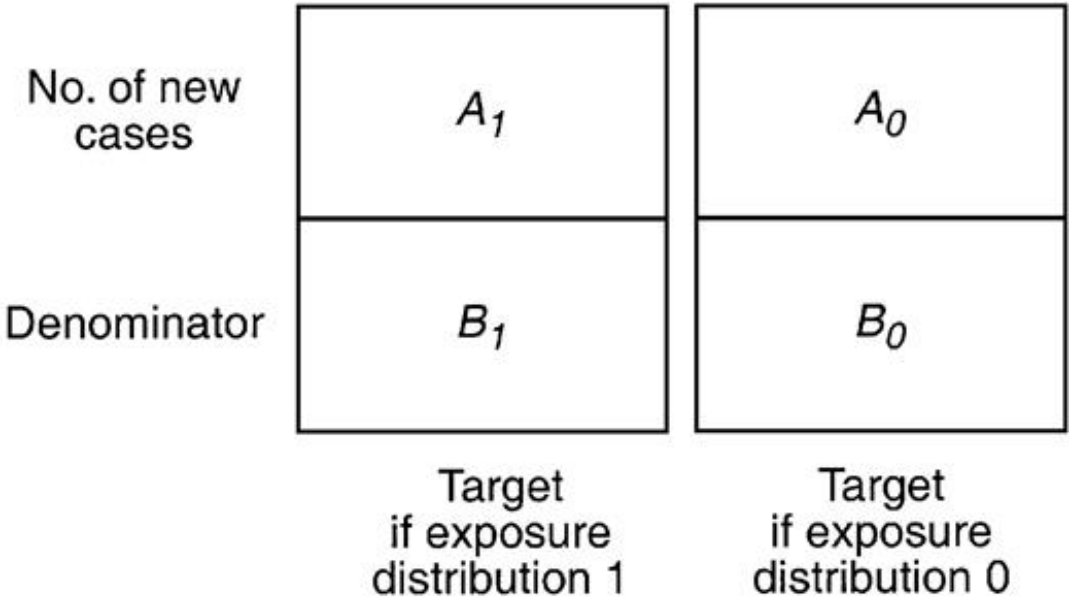
- If B = people at risk at the *beginning of the period* [2] = **incidence proportion, average risk**.
- If B = person-time at risk during the period, R = **person-time incidence rate**.
- If B = people who do *NOT* get disease by the end of the period, R = **incidence odds.**

[1] *International Journal of Epidemiology, Volume 31, Issue 2, April 2002, Pages 422–429, https://doi.org/10.1093/intjepid/31.2.422*

[2] *And all individuals are followed throughout the etiologic time period.*

Define the causal effect

Define the counter-to-fact condition and outcome in the target:



Only possible to, at best, observe one of these conditions.

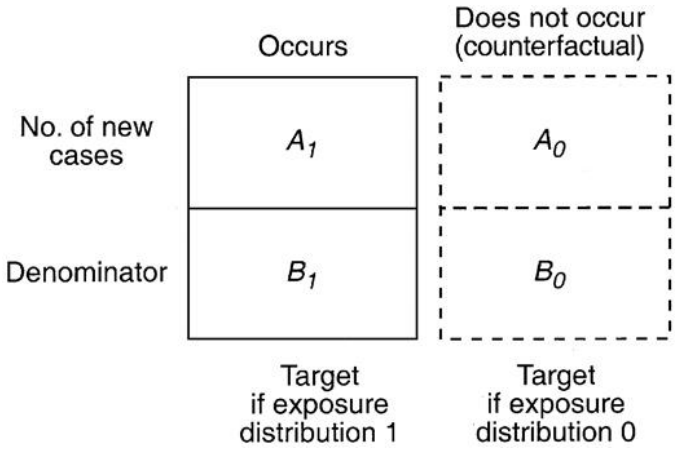
Define the causal effect

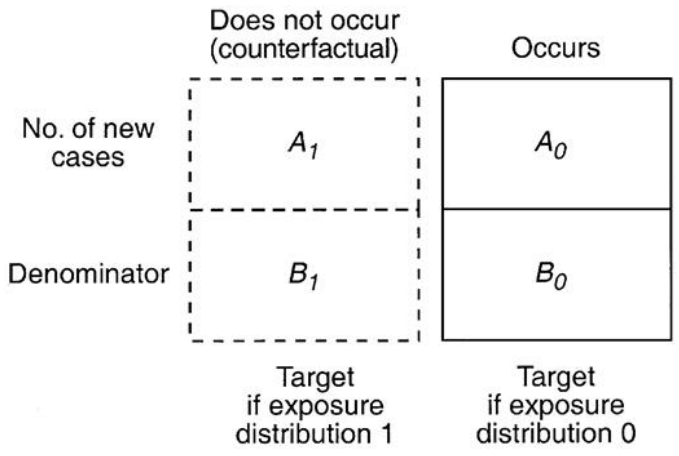
Let $R_1 = A_1/B_1$ and $R_0 = A_0/B_0$

- $R_1 - R_0$ is the causal **difference measure** and
- R_1/R_0 is the **causal ratio measure**
- Both of these are causal contrasts (measures of effect).
- Here, the only possible reason for a difference between R_1 and R_0 is due to **exposure** because we are contrasting the **exact same people over the exact same time period**.
- But, we **cannot observe the causal contrast** because we cannot observe both conditions.

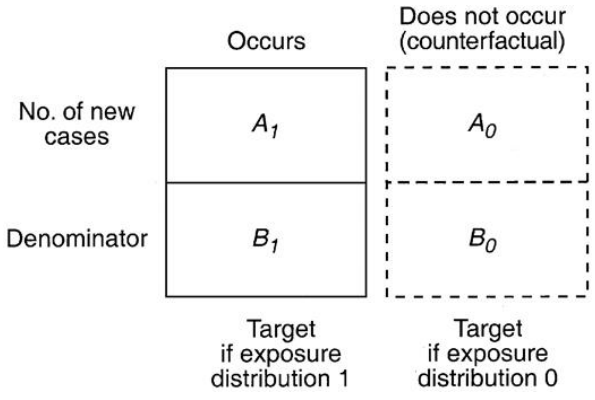
We require a substitute with observable information!

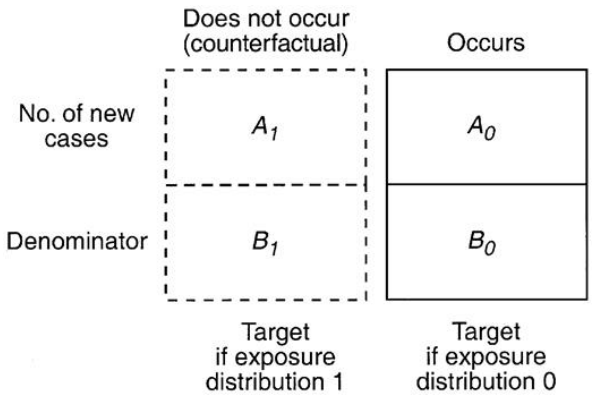
Can't observe the counterfactual

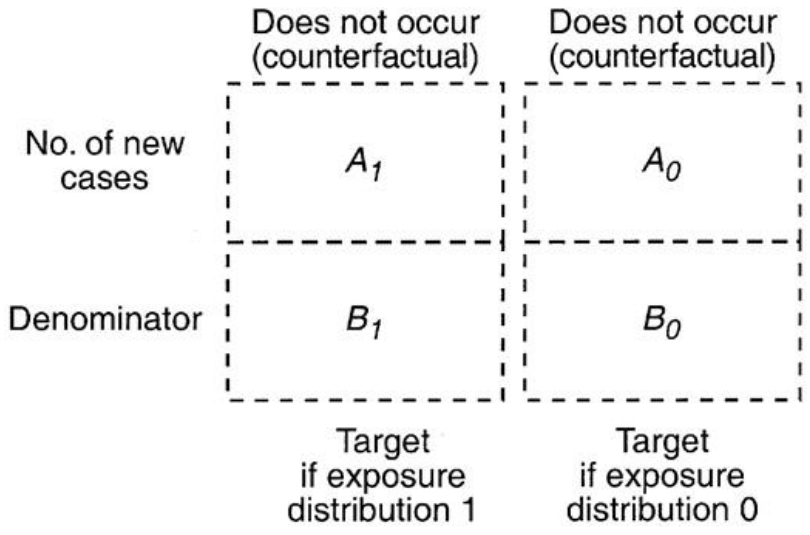




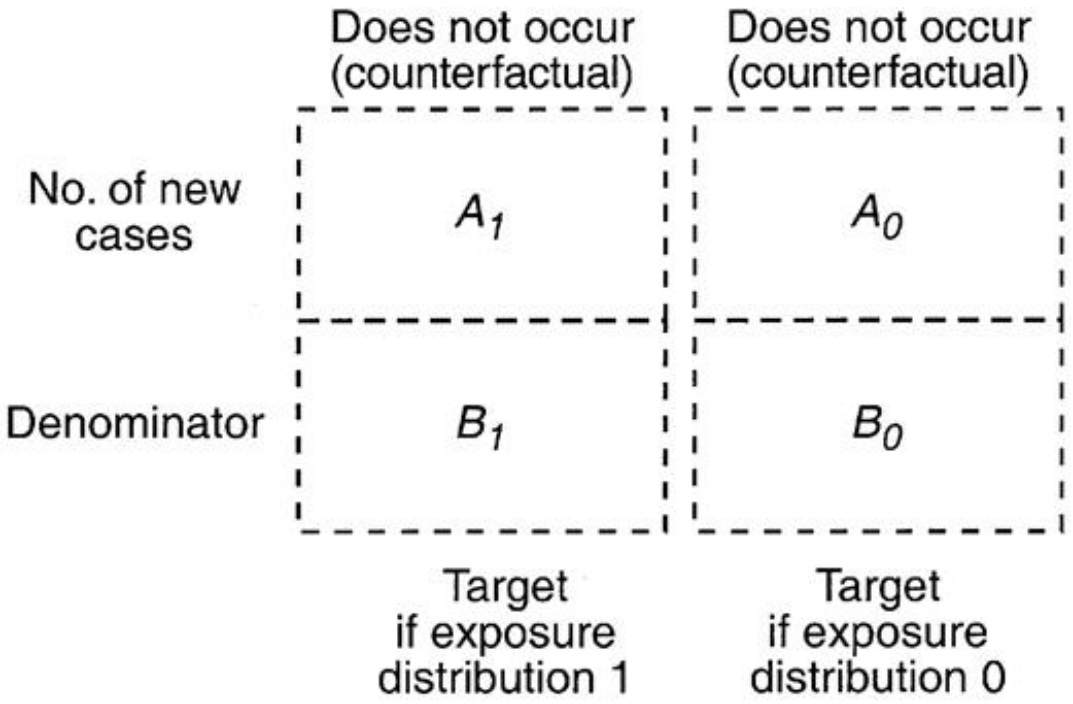
Can't observe the counterfactual





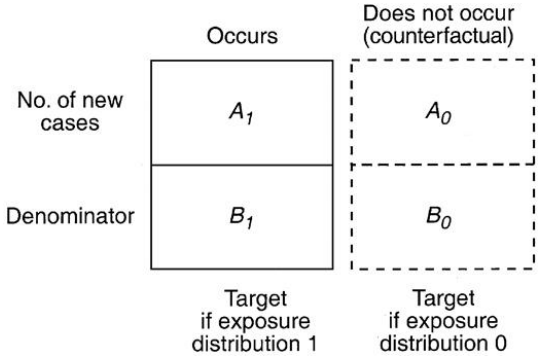


Can't observe the counterfactual



> "*Both R_1 and R_0 are counterfactual disease frequencies,both are hypothetical alternatives to the actual disease frequency that occurs under the actual exposure distribution (which is neither exposure distribution 1 nor 0), and therefore neither R1 nor R0 can occur and be observed.*" Maldonado & Greenland 2002

Can't observe the counterfactual



The impossibility of observing both halves of the causal contrast leads to the idea of substitute populations.

These are often:

- Different people observed during the same etiologic time period.
- The same people observed over two different time periods (case-crossover design).

Maldonado & Greenland 2002

Defining the substitute population

In a substitute population under **exposure distribution** 1, let

- C_1 be the name for the numerator of the disease-frequency measure,
- D_1 be the denominator (number of people or amount of person-time at risk).

In a substitute under **exposure distribution** 0, let

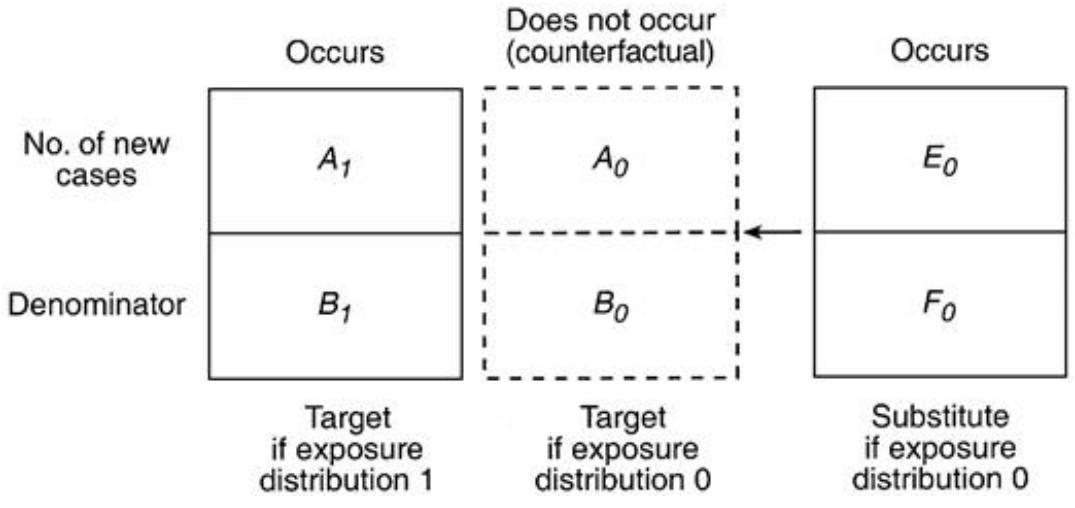
- E_0 be the numerator,
- F_0 be the denominator.

Exposure = 1, C_1/D_1

Exposure = 0, E_0/F_0

Defining the substitute population

Target experiences **exposure distribution 1**

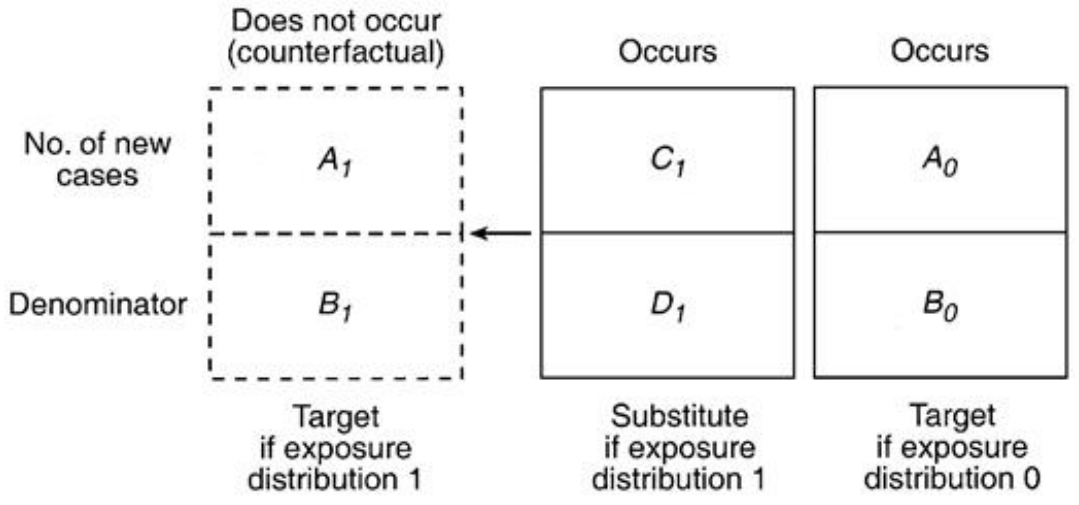


$$RR_{association} = \frac{R_1}{\text{Substitute for } R_0} = \frac{R_1}{E_0/F_0} = \frac{A_1/B_1}{E_0/F_0}$$

Maldonado & Greenland 2002

Defining the substitute population

Target experiences **exposure distribution 0**

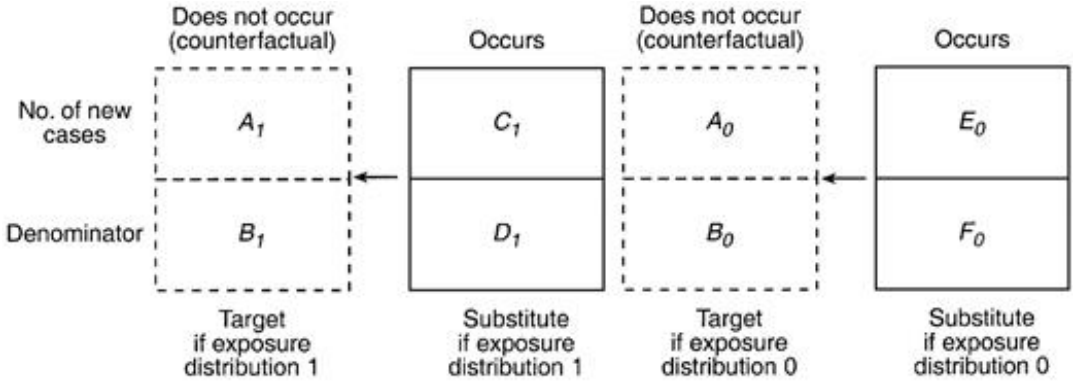


$$RR_{association} = \frac{\text{Substitute for } R_1}{R_0} = \frac{C_1/D_1}{R_0} = \frac{C_1/D_1}{A_0/B_0}$$

Maldonado & Greenland 2002

Defining the substitute population

Target experiences neither **exposure distribution 1 or 0**



$$RR_{association} = \frac{\text{Substitute for } R_1}{\text{Substitute for } R_0} = \frac{C_1 / D_1}{E_0 / F_0}$$

Maldonado & Greenland 2002

Other Notation Used

- Greenland employs a probabilistic model of disease such that each individual i has a risk r_{1_i} of disease when $E = 1$ and a risk of r_{0_i} when $E = 0$.
- Survival probabilities: $S_{1_i} = 1 - r_{1_i}$ and $S_{0_i} = 1 - r_{0_i}$
- Odds: $w_{1_i} = r1_i / s_{1_i}$ and $w_{0_i} = r_{0_i} / s_{0_i}$.
    - Only defined when survival probabilities are not equal to zero.

Notation Used

- The effect of exposure on the risk of an individual can be measured in terms of the risk difference $r_{1_i} - r_{0_i}$, risk ratio r_{1_i}/r_{0_i}, or the risk-odds ratio w_{1_i}/w_{0_i}

- The ratios will be undefined if the risk in the exposed group is 0 and

- The risk-odds ratio will be undefined if either survival probability is 0

Other Notation Used

In a cohort with N_1 $E+$ individuals and N_0 $E-$ individuals:

	E+	E-
D+	$A = \Sigma_1 r_{1i}$	$B = \Sigma_0 r_{0i}$
D-	$C = \Sigma_1 s_{1i}$	$D = \Sigma_0 s_{0i}$
Total	N_1	N_0

- Incidence proportions: A/N_1 and B/N_0, interpretable as average risks in their respected groups
- Incidence odds: A/C and B/D, interpretable as ratios of the average risk to the average survival probabilities.

Defining the counterfactual

- Assuming **no confounding** [1]
- Had the **exposure** been absent from the $E+$ group,
- The average risk would have been **the same** among the sub-cohorts that were in fact exposed and unexposed.

$$\left(\frac{\sum_1 r_{0i}}{N_1}\right) = \left(\frac{\sum_0 r_{0i}}{N_0}\right)$$

[1] More on confounding next lecture.

Risk Difference

Thus: The risk difference is interpretable as both:

1) the **absolute change in the average risk** of the exposed sub-cohort produced by exposure,

$$\left(\frac{\sum_1 r_{1i}}{N_1}\right) - \left(\frac{\sum_1 r_{0i}}{N_1}\right)$$

2) the **average absolute change in risk** produced by exposure among exposure individuals

$$\left(\frac{\sum_1 (r_{1i} - r_{0i})}{N_1}\right)$$

Expressions 1 and 2, on: Interpretation and choice of effect measures in epidemiologic analyses. S. Greenland (1987)

Risk Ratio

The incidence proportion ratio is given by

$$\left(\frac{A}{N_1}\right) / \left(\frac{B}{N_0}\right)$$

The risk ratio is interpretable as :

1) the **proportionate change in the average risk** of the exposed subcohort produced by exposure,

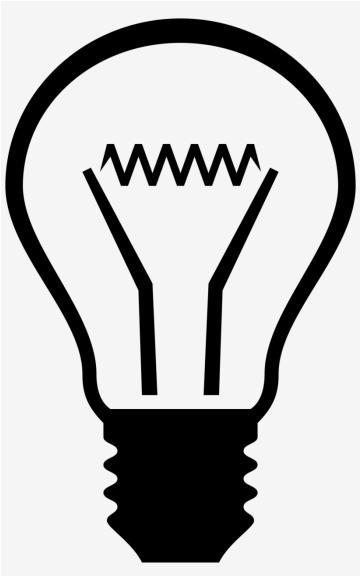
$$\left(\frac{\sum_1 r_{1i}}{N_1}\right) / \left(\frac{\sum_0 r_{0i}}{N_0}\right) = \left(\frac{\sum_1 r_{1i}}{N_1}\right) / \left(\frac{\sum_1 r_{0i}}{N_1}\right)$$

It **is not interpretable** as the average proportionate change in risk produced by exposure among exposed individuals:

$$\left(\frac{\sum_1 (r_{1i}/r_{0i})}{N_1}\right)$$

Expressions 3 and 4, on: Interpretation and choice of effect measures in epidemiologic analyses. S. Greenland (1987)

Incidence Proportion Ratio



However, **if the individual risk ratios are all equal** then the ratio of the average risks across exposure will be equal to the average of the individual risk ratios

A note on Risk Ratios

Risk Differences have a symmetric range [-1, 1]

But the risk Ratios have an ***asymmetric*** range:

- From 0 to 1, below the null
- From 1 to infinity above the null

This presents a challenge in the interpretation...

What's more "impressive" `(RR = 2)` **or** `(RR= 0.2)` **?** ...

What's more "impressive" a (RR = 2) or a (RR= 0.2) ?

Two ways to find out:

Simple: the reciprocal of the value below the null = 1/0.2 = 5, since 5 > 2, then

- a RR of 0.2 is of larger magnitude (further away form the null) than a RR = 2

Elaborated: Take the absolute values of the natural logarithm (log or ln) of each value:

- log(2) = 0.693
- log(0.2) = -1.609

In absolute terms, |log(0.2)| **>** |log(2)| = **|-1.609| > |0.693|**

- Regression models for ratio measures generally operate on the log-scale (that's why we exponentiate to provide estimates and graph on the log scale).
- Note: Recall that on the log10 or ln scale, the null for a ratio measure is 0, not 1 – because log(1) = ln(1) = 0.
- Try it with 3 and 0.3, and with 5 and 0.5 and see what happens! :)

Incidence Odds Ratio

The incidence odds ratio is given by:

$$\left(\frac{A}{C}\right) / \left(\frac{B}{D}\right)$$

Thus, the odds ratio is interpretable as :

1) the proportionate change in the incidence odds in the exposed subcohort produced by exposure,

$$\left(\frac{\sum_1 r_{1i}}{\sum_1 s_{1i}}\right) / \left(\frac{\sum_0 r_{0i}}{\sum_0 s_{0i}}\right)$$

$$\left(\frac{\sum_1 r_{1i}}{\sum_1 s_{1i}}\right) / \left(\frac{\sum_1 r_{0i}}{\sum_1 s_{0i}}\right)$$

It **is not interpretable** as the proportionate change in the average odds in the exposed produced by exposure:

Incidence Odds Ratio

- Furthermore,neither of the last two expressions is equivalent to the average of the individual odds ratios among the exposed

$$\left(\frac{\sum_1 (w_{1_i} / w_{0_i})}{N_1} \right)$$

- The incidence odds ratio **(that we calculate)** lacks any simple interpretation in terms of exposure effect on the average risk or odds, or average exposure effect on individual risk or odds.

Incidence Odds Ratio

The incidence odds do not equal the simple averages of the risk odds:

This severely handicaps the interpretability of measures based on the incidence odds.

- It is not a measure of average causal effect (the RR and RD are) (Greenland 1987)
- Cannot be relied upon to reveal confounding (Greenland et al., 1999)

Incidence Odds Ratio

- If the individual ORs are all equal (which is the assumption made by a logistic model), then the ratio of the average odds will equal the average of the individual odds ratios.

But, the incidence odds ratio will need not equal that value!

Incidence Odds Ratio

- For example: Define a population where 10% of people have $r_{1i} = 0.60$ and
- $r_{0i} = 0.20$ and 90% of the people have $r_{1i} = 0.035$ and $r_{0i} = 0.006$
- Here, the individual ORs = 6.0 for every individual
    - the average of the individual ORs = 6.0
- Also, the ratio of the average odds equals 6.0 as well

But, the incidence odds ratio is equal to 3.9

Want to give it a try and calculate it?

Incidence Odds Ratio

- Because of this fact, the crude odds ratio can be smaller than any of the stratum-specific odds ratios, even if confounding is entirely absent

- This paradoxical behaviour will not occur with the risk difference or the risk ratio

- Unless, equal to the null (when OR = RR = 1) the OR will ***almost always*** be further away from the null than RRs.
    - 2nd exception is when **OR = RR = 0**, as would occur when risk in the exposed is 0 (zero), and risk in the unexposed is for example 0.6. In this case, the risk ratio is 0 and the OR is odds(0)/odds(0.6), which is also 0.

More on Epidemiology by design by Daniel Westreich

Odds difference?

⚠ !! Nope, nope, nope!!!!!! ⚠ !!

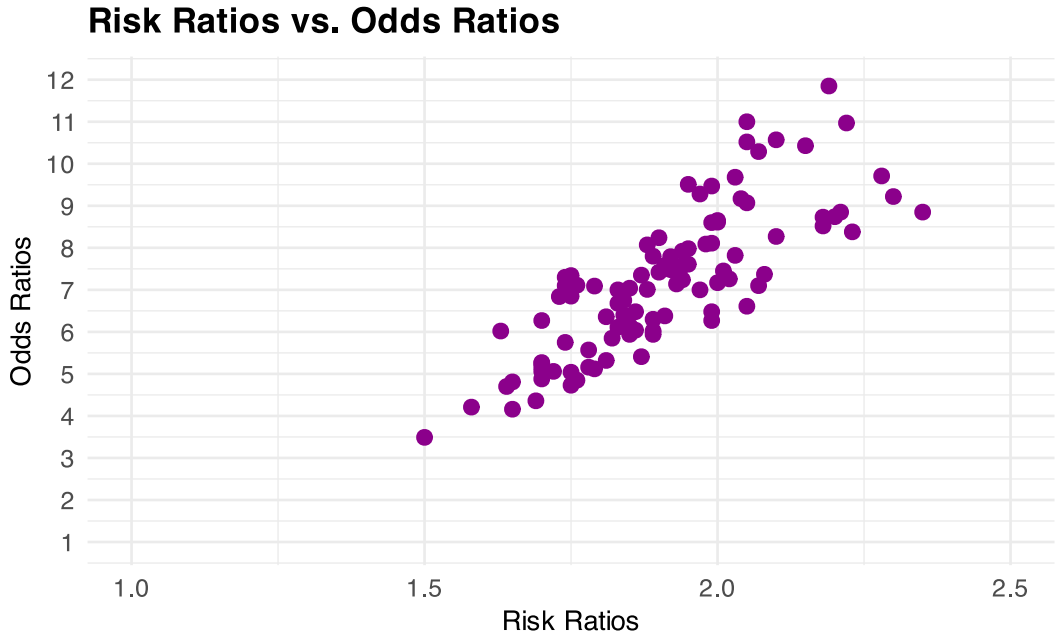
✕ !! We never do this. !! ✕

Some simulated Examples - Common Outcome

```
set.seed(7042025)
yea.dat1 <- function(n) {
          E <- rbinom(n,1,0.55) #parameters for E
          Y <- ifelse(E==1, rbinom(n,1,0.85), rbinom(n, 1, 0.45)) #parameters for Y

    return(data.frame(E=E,Y=Y)) #ask to return a data set with those parameters
}
sim100 <- lapply(1:100,FUN=function(x) yea.dat1(400))
RRs <- sapply(sim100,FUN=function(x) {
   results <- logbin(Y ~ E , data=x)$coef
   return(round(exp(results[names(results)=="E"]),2))
 })

ORs <- sapply(sim100,FUN=function(x) {
   results <- glm(Y ~ E , family="binomial", data=x)$coef
   return(round(exp(results[names(results)=="E"]),2))
 })

sim_RRs1<-round(quantile(RRs,probs = c(0.05,0.5,0.95)),2)
sim_ORs1<-round(quantile(ORs,probs = c(0.05,0.5,0.95)),2)
```

Some simulated Examples - Common Outcome

```
##              5%  50%   95%
## sim_RRs1 1.65 1.90  2.22
## sim_ORs1 4.73 7.12 10.52
```



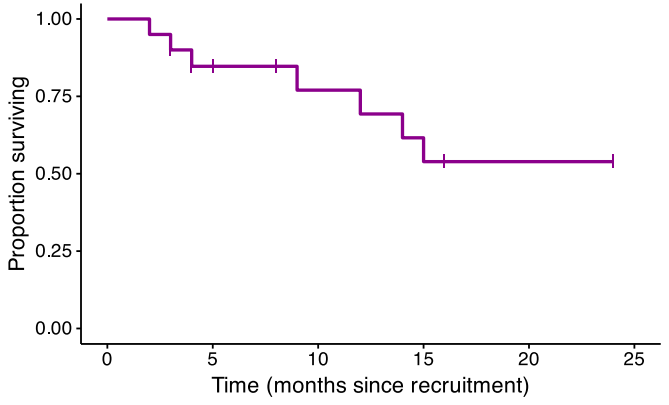
Key Takeaways

Only incidence difference and incidence ratios possess direct interpretations as measures of impact on average risk/hazard

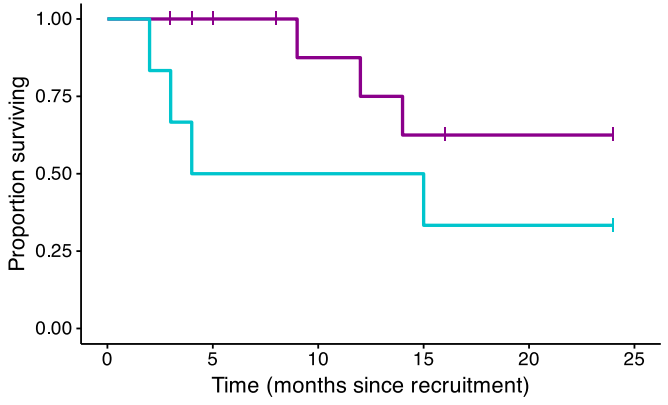
Consequently, odds ratios are useful only when they serve as incidence ratio estimates

Incidence Rate Differences and Ratio

From everyone's time...



... To contrast outcomes and time



Incidence Rate Differences and Ratio

- Incidence rate of outcome Y when X=1 is $IR(Y = 1|X = 1) = IR(Y|X = 1)$
- Incidence rate of outcome Y when X=0 is $IR(Y = 1|X = 0) = IR(Y|X = 0)$
- **Incidence rate difference:**
  - Range: -∞ to ∞
  - Null value: 0
- **Incidence rate ratio**:
  - Range: 0 to ∞
  - Null value: 1

Relationship among Risk, Odds, and Incidence Rates

Question: If we calculated the risk ratio, odds ratio and rate ratio, which will be closest to the null? Furthest from the null?

Notation:

- R = Incidence Proportion ("Risk"),
- $S = 1 - R$ ("Survival Proportion"),
- I = Incidence Rate
- T = interval length
- i = 1 if exposed; i= 0 if unexposed

Relationship among Risk, Odds, and Incidence Rates

Relations among relative risks

- In a closed population where the population at risk declines only slightly over the interval (implying that R must be small and S is close to 1): $R \cong I\Delta T \cong R/S$

This implies:

$$\left(\frac{R_1}{R_0}\right) \cong \left(\frac{I_1\Delta T_1}{I_0\Delta T_0}\right) \cong \left(\frac{I_1}{I_0}\right) \cong \left(\frac{R_/S_1}{R_0/S_0}\right)$$

[Numerators] Holds if R_1 and R_0 are small enough so that S_1 and S_0 are close to 1

[Two Denominators on the right] Holds if exposure only has negligible effects on the person-time at risk

Relationship among relative risks

- If exposure causes the outcome, then R1>R0 and S1<S0.

$$1 < \left(\frac{R_1}{R_0}\right) < (\left(\frac{R_1}{R_0}\right) \times \left(\frac{S_0}{S_1}\right)) = \left(\frac{R_1/S_1}{R_0/S_0}\right)$$

- If exposure prevents the outcome, $R1 < R0$ and $S1 > S0$, such that:
  $1 > \left(\frac{R_1}{R_0}\right) > (\left(\frac{R_1}{R_0}\right) \times \left(\frac{S_0}{S_1}\right)) = \left(\frac{R_1/S_1}{R_0/S_0}\right)$

In words: **The odds ratio is further from the null than the risk ratio**

Relationship among relative risks

- Now, if exposure is harmful $(R1 > R0)$ then we would ordinarily expect exposure to reduce the person-time at risk $(T1 < T0)$,
- and if exposure is preventive $(R1 < R0)$ then we expect exposure to increase the person-time at risk $(T1 > T0)$.
- Thus, when exposure is harmful:

$$1 < \left(\frac{R_1}{R_0}\right) \cong \left(\frac{I_1 \Delta T_1}{I_0 \Delta T_0}\right) < \left(\frac{I_1}{I_0}\right)$$

Relationship among relative risks

And when exposure is preventive:

$$1 > \left(\frac{R_1}{R_0}\right) \cong \left(\frac{I_1 \Delta T_1}{I_0 \Delta T_0}\right) > \left(\frac{I_1}{I_0}\right)$$

- In words: We would **ordinarily** expect the risk ratio to be closer to the null than the rate ratio. Under further conditions, the rate ratio will be closer to the null than the odds ratio (Greenland and Thomas, 1982)

Relationship among relative risks

Thus, we usually expect:

- Risk ratio nearest to the null
  - *implicitly* suggesting all events occur at the end of follow up
- Odds ratio furthest from the null
  - *implicitly* suggesting all events occur at the beginning of follow up
- Rate ratio somewhere in between
  - allows event to occur at any point in time

1 < Risk Ratio < Rate Ratio < Odds Ratio

[1] More on Epidemiology by design by Daniel Westreich

Prevalence Ratios

Recall that the prevalence odds is equal to the incidence rate multiplied by the average duration in a stationary, closed population. This implies:

$$POR = \left(\frac{PO_1}{PO_0}\right) = \left(\frac{I_1\overline{D_1}}{I_0\overline{D_0}}\right) = \left(\frac{I_1}{I_0}\right)$$

if the average duration of disease is unaffected by exposure.

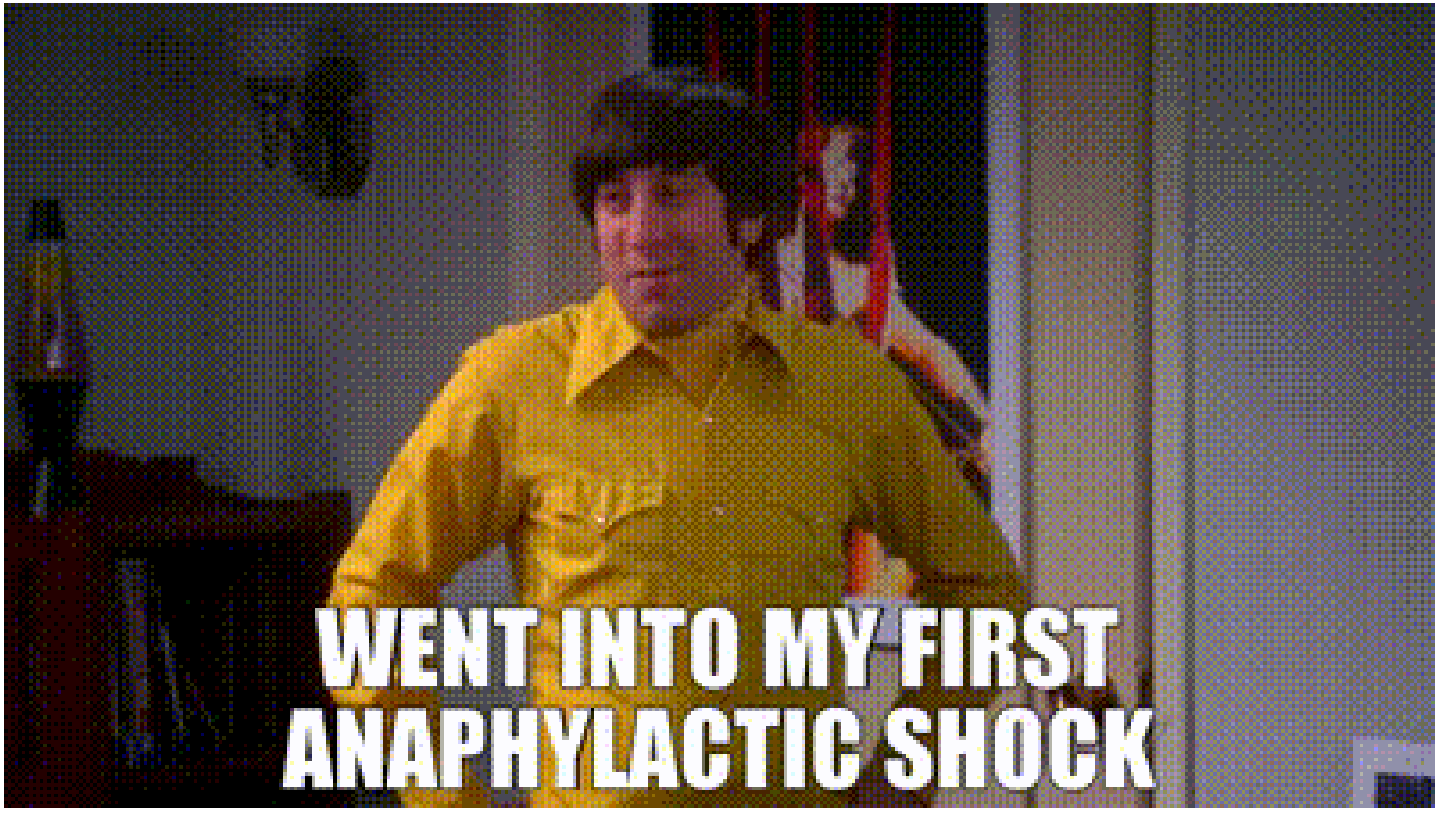
Prevalence Ratios

- If prevalence is low, then the prevalence ratio is approximately equal to the incidence rate multiplied by average duration. This implies:

$$PR = \left(\frac{P_1}{P_0}\right) \cong \left(\frac{I_1\overline{D_1}}{I_0\overline{D_0}}\right) = \left(\frac{I_1}{I_0}\right)$$

if the average duration of disease is unaffected by exposure

...Still breathing? :)



Key Points

> - *"There was a decreasing incidence of allergic reactions overall, in both sexes and all age subgroups. The decrease in frequency did not depend on age or sex."*
> - *"Our multivariable logistic regression analysis showed that GPI IIb/IIIa, LMWH, previous PCI, contrast dose, and x-ray dose were independent predictors of allergic reaction occurrence".*

[1] Allergic reactions during coronary angiography or PCI in Poland: Occurrence, trends, and long-term perspective based on the Polish ORPKI registry. Polish Heart Journal (Kardiologia Polska)Vol 83, No 3 (2025)

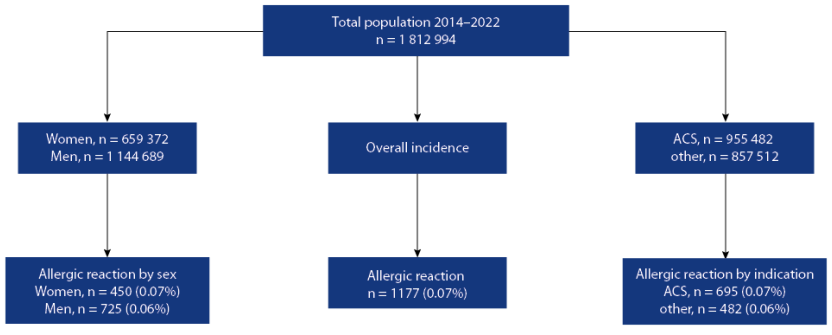
Illustrated Example:

Table 3. Clinical and periprocedural factors affecting incidence allergic reaction

Variables	Univariable	
	OR (95% CI)	*P*-value
Male sex	0.93 (0.83–1.04)	0.21
Age, years	1 (0.99–1)	0.24
Weight, kg	1 (1–1)	0.49
Diabetes	0.87 (0.75–1)	0.06
Previous stroke	1.14 (0.83–1.57)	0.42
Previous myocardial infarction	1.27 (1.12–1.44)	<0.001
Previous PCI	1.37 (1.22–1.55)	<0.001
Previous CABG	0.89 (0.68–1.16)	0.37
Smoking status (active)	0.8 (0.68–0.94)	0.01
Arterial hypertension	0.96 (0.85–1.09)	0.56
Kidney disease	0.68 (0.5–0.92)	0.01
COPD	0.46 (0.27–0.78)	0.003
ACS (yes/no)	1.29 (1.15–1.45)	<0.001
UFH during angiogram	0.67 (0.52–0.87)	0.003
LMWH during angiogram	0.53 (0.2–1.41)	0.21
Cardiac arrest at baseline	1.24 (0.62–2.48)	0.55
Direct transport to cathlab	0.64 (0.41–0.99)	0.045
UFH	0.34 (0.3–0.39)	<0.001
LMWH	49.18 (43.85–55.15)	<0.001

[1] Allergic reactions during coronary angiography or PCI in Poland (2025)

Illustrated Example... Let's reproduce the analysis



	Outcome		
	Anaphylaxis	Non.Anaphylaxis	TOTAL
Male	687	1144002	1144689
Female	462	688910	659372
Total	1149	1832912	1804061

Allergic reactions during coronary angiography or PCI in Poland. Polish Heart Journal (Kardiologia Polska)Vol 83, No 3 (2025)

Illustrated Example... Let's reproduce the analysis from Figure 1

```
pcidat1<-c(687, 1144002, 462, 658943)

pciOR1<- epi.2by2(pcidat1, method = "cross.sectional")
pciOR1$tab
```

```
##               Outcome +    Outcome -      Total                Prev risk *
## Exposed +           687      1144002    1144689       0.06 (0.06 to 0.06)
## Exposed -           462       658943     659405       0.07 (0.06 to 0.08)
## Total              1149      1802945    1804094       0.06 (0.06 to 0.07)
```

"No gender differences" **(???)**

```
round(pciOR1$massoc.detail$OR.strata.wald, 2)
```

```
##    est lower upper
## 1 0.86  0.76  0.96
```

Illustrated Example... Let's reproduce the analysis

```
pciOR1<- epi.2by2(pcidat1, method = "cross.sectional")
pciOR1
```

```
##               Outcome +    Outcome -      Total                 Prev risk *
## Exposed +           687      1144002    1144689       0.06 (0.06 to 0.06)
## Exposed -           462       658943     659405       0.07 (0.06 to 0.08)
## Total              1149      1802945    1804094       0.06 (0.06 to 0.07)
##
## Point estimates and 95% CIs:
## -------------------------------------------------------------------
## Prev risk ratio                                 0.86 (0.76, 0.96)
## Prev odds ratio                                 0.86 (0.76, 0.96)
## Attrib prev in the exposed *                    -0.01 (-0.02, -0.00)
## Attrib fraction in the exposed (%)            -16.74 (-31.35, -3.76)
## Attrib prev in the population *                 -0.01 (-0.01, 0.00)
## Attrib fraction in the population (%)         -10.01 (-18.04, -2.52)
## -------------------------------------------------------------------
## Uncorrected chi2 test that OR = 1: chi2(1) = 6.635 Pr>chi2 = 0.010
## Fisher exact test that OR = 1: Pr>chi2 = 0.011
##  Wald confidence limits
##  CI: confidence interval
##  * Outcomes per 100 population units
```

How do I obtain the χ^2 ?

Illustrated Example... What's in the Table 3?

```
pcidat1a<-c(725, 1144689, 450, 659372)
pciOR1a<- epi.2by2(pcidat1a, method = "cross.sectional")
pciOR1a
```

```
##               Outcome +    Outcome -      Total                 Prev risk *
## Exposed +           725      1144689    1145414       0.06 (0.06 to 0.07)
## Exposed -           450       659372     659822       0.07 (0.06 to 0.07)
## Total              1175      1804061    1805236       0.07 (0.06 to 0.07)
##
## Point estimates and 95% CIs:
## -------------------------------------------------------------------
## Prev risk ratio                                 0.93 (0.83, 1.04)
## Prev odds ratio                                 0.93 (0.83, 1.04)
## Attrib prev in the exposed *                    -0.00 (-0.01, 0.00)
## Attrib fraction in the exposed (%)             -7.75 (-21.19, 4.20)
## Attrib prev in the population *                 -0.00 (-0.01, 0.00)
## Attrib fraction in the population (%)          -4.78 (-12.67, 2.55)
## -------------------------------------------------------------------
## Uncorrected chi2 test that OR = 1: chi2(1) = 1.548 Pr>chi2 = 0.213
## Fisher exact test that OR = 1: Pr>chi2 = 0.214
##  Wald confidence limits
##  CI: confidence interval
##  * Outcomes per 100 population units
```

What's clear and what's not here?

Results

of significance of the difference in years (Table 2). Among clinical and pharmacological factors, the risk of allergic reactions was increased by a history of myocardial infarction, a history of coronary artery by-pass grafting, a diagnosis of ST-segment elevation myocardial infarction (STEMI), a diagnosis of ACS, the use of IIb/IIIa glycoprotein inhibitor (GPI IIb/IIIa), low molecular weight heparin (LMWH), and earlier PCI. The risk of allergic reaction was reduced by active smoking, renal disease, chronic obstructive pulmonary disease, UFH, and performing coronary angiography only (Table 3).

Allergic reactions during coronary angiography or PCI in Poland. Polish Heart Journal (Kardiologia Polska)Vol 83, No 3 (2025)

Methods

Statistical analysis

In this study, nominal variables were represented as percentages along with their respective counts. The temporal variation in the proportion of allergic reactions was analyzed using the Cochran–Armitage test for trend. The logistic regression model was employed to identify the predictors of periprocedural allergic reactions. All potential demographic, baseline, and procedural characteristics were incorporated into a multivariable model if the *P*-value in a univariable model was less than 0.2 or if they were considered to have clinical significance. The final model was derived by minimizing the Akaike information criterion. The robustness of the model was validated using bootstrap resampling. To evaluate the presence of multicollinearity, variance inflation factors were analyzed. The results were expressed as odds ratios with 2-sided 95% confidence intervals. The entire statistical analysis was conducted using R software (version 4.3.1)

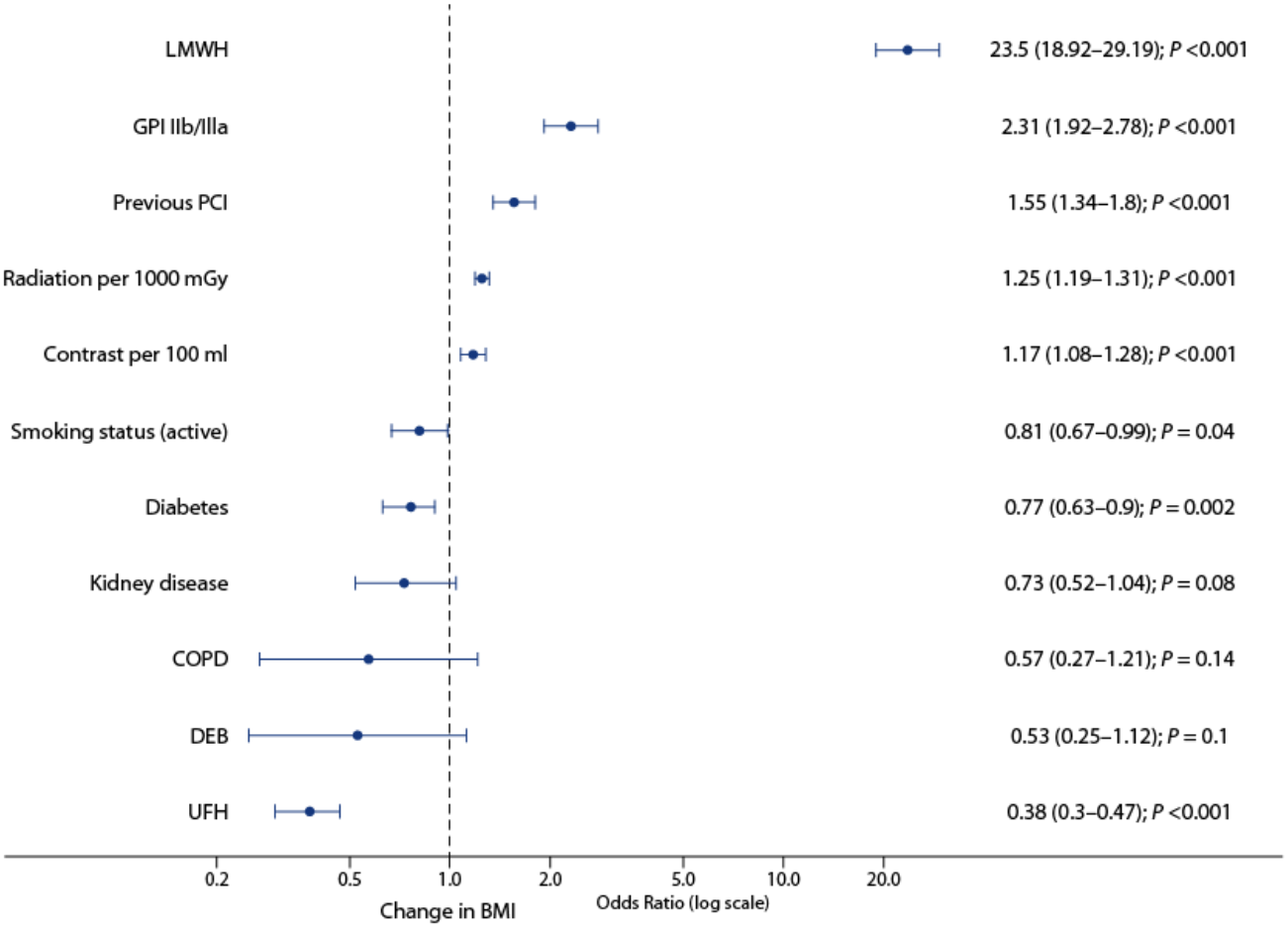
Illustrated Example

Table 3. Clinical and periprocedural factors affecting incidence allergic reaction

Variables	Univariable	
	OR (95% CI)	*P*-value
Male sex	0.93 (0.83–1.04)	0.21
Age, years	1 (0.99–1)	0.24
Weight, kg	1 (1–1)	0.49
Diabetes	0.87 (0.75–1)	0.06
Previous stroke	1.14 (0.83–1.57)	0.42
Previous myocardial infarction	1.27 (1.12–1.44)	<0.001
Previous PCI	1.37 (1.22–1.55)	<0.001
Previous CABG	0.89 (0.68–1.16)	0.37
Smoking status (active)	0.8 (0.68–0.94)	0.01
Arterial hypertension	0.96 (0.85–1.09)	0.56
Kidney disease	0.68 (0.5–0.92)	0.01
COPD	0.46 (0.27–0.78)	0.003
ACS (yes/no)	1.29 (1.15–1.45)	<0.001
UFH during angiogram	0.67 (0.52–0.87)	0.003
LMWH during angiogram	0.53 (0.2–1.41)	0.21
Cardiac arrest at baseline	1.24 (0.62–2.48)	0.55
Direct transport to cathlab	0.64 (0.41–0.99)	0.045
UFH	0.34 (0.3–0.39)	<0.001
LMWH	49.18 (43.85–55.15)	<0.001

[1] Allergic reactions during coronary angiography or PCI in Poland. (2025)

What's clear and what's not here?



What would be the actual interpretation of these ORs?

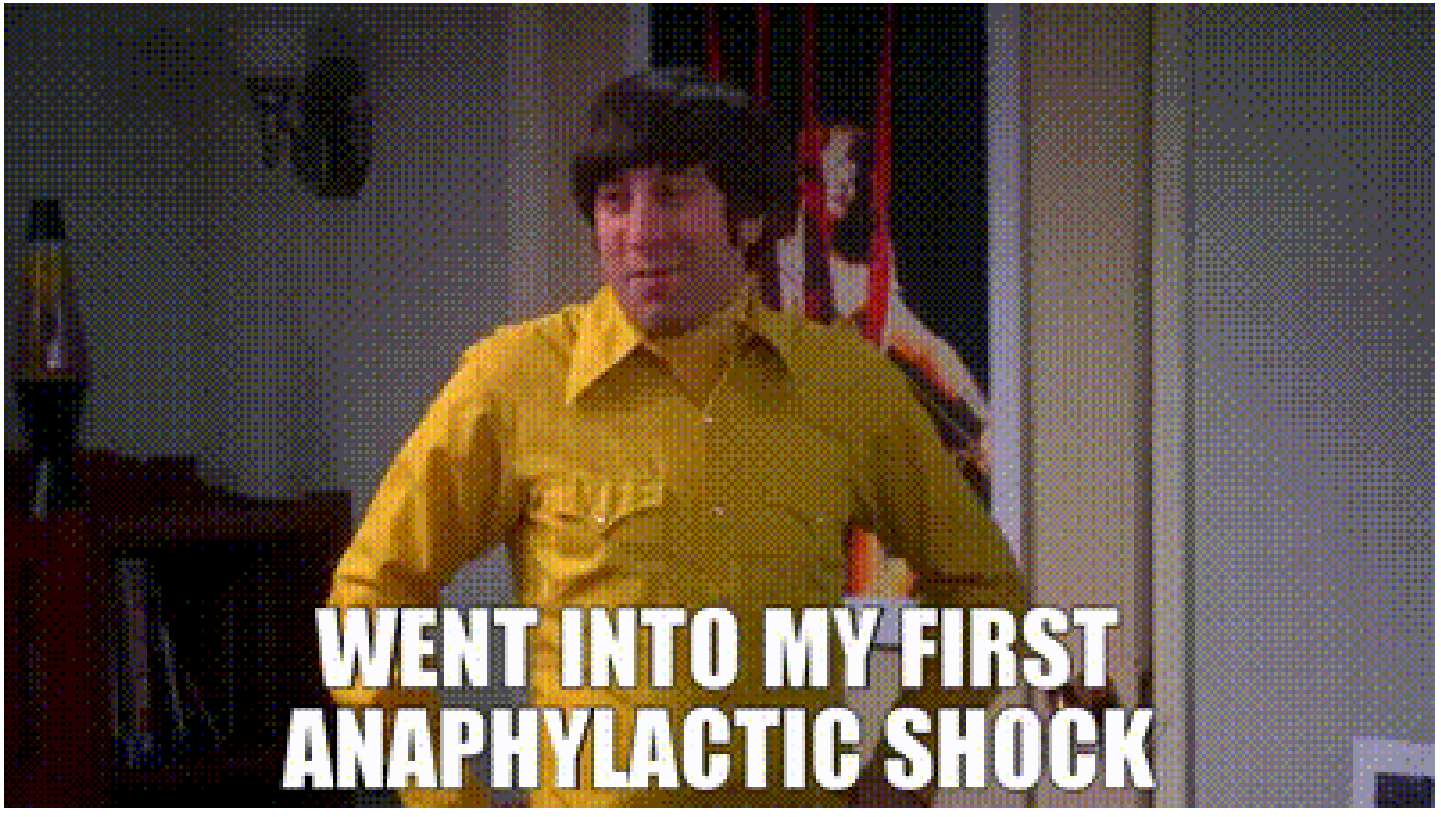
Illustrated Example

> *"When assessing the incidence of allergic reactions ..., there was a decrease in the frequency of allergic reactions regardless of acute coronary syndromes (ACS) (Table 1)."*
> *"...the risk of allergic reactions was increased by... a diagnosis of ACS"*

```
pcidat2<-c(695, 955482, 482, 857512)
pciOR2<- epi.2by2(pcidat2, method = "cross.sectional")
pciOR2
```

```
##               Outcome +    Outcome -      Total                Prev risk *
## Exposed +           695       955482     956177       0.07 (0.07 to 0.08)
## Exposed -           482       857512     857994       0.06 (0.05 to 0.06)
## Total              1177      1812994    1814171       0.06 (0.06 to 0.07)
##
## Point estimates and 95% CIs:
## -------------------------------------------------------------------
## Prev risk ratio                                 1.29 (1.15, 1.45)
## Prev odds ratio                                 1.29 (1.15, 1.45)
## Attrib prev in the exposed *                    0.02 (0.01, 0.02)
## Attrib fraction in the exposed (%)             22.71 (13.19, 31.19)
## Attrib prev in the population *                 0.01 (0.00, 0.01)
## Attrib fraction in the population (%)          13.41 (7.26, 19.15)
## -------------------------------------------------------------------
## Uncorrected chi2 test that OR = 1: chi2(1) = 19.007 Pr>chi2 = <0.001
## Fisher exact test that OR = 1: Pr>chi2 = <0.001
##  Wald confidence limits
##  CI: confidence interval
##  * Outcomes per 100 population units
```

What did the authors do with the covariates?



When to use which measure of Association

- Research Question
- Public Health Relevance
- Intervention (Design of/ Intervenable exposure?)

- Study design

- Contrasts requires assignment to group, which requires measurement of group membership For example, measurement of exposure
    - What will happen if the exposure is measured poorly?
- Since risks are only well-defined within a specific time-period, state that time-period.

Things to consider when measuring associations

- Be aware that in real life, we encounter:
  - Random Error
  - Systematic error
  - Competing Risks, Confounding, Selection Bias, Measurement Error
  - Methods related limitations
  - Clinical vs Statistical hurdles

Contrast (Association vs Impact)

Most contrasts can be used to assess either:

- Association (which is agnostic on the question of causality)
- Impact (which is causal). The risk difference, for example.

Certain measures, however, are implicitly causal and (probably) shouldn't be used to merely describe an association.

- Among these latter measures are **number needed to treat, and attributable contrasts**.

Population Attributable Fraction (PAF)

A population attributable fraction (PAF) can be thought of as

- "*The proportion of disease burden among the total population which is caused by the exposure*".
- That definition is explicitly causal.
- PAF is implicitly causal.
- “Attribution” implies cause, though we can argue over usage.

Calculated as $(P(Y) - P(Y^{x=0}))/P(Y)$ where,

- $P(Y)$ is the risk of the outcome in the whole population,
- $P(Y|X = 0)$ is the risk of the outcome in the unexposed.

Note that because most outcomes are caused by more than one thing, the sum of PAFs can be (and often are) greater than 100%. Epidemiology by design by Daniel Westreich

Example:

- Exposure to TB is a necessary cause of active TB: by definition.
  - So from the above, $P(Y_{TB}exposure = 0) = 0$, and so PAF = 100%.
- But not everyone exposed to TB develops active TB. There are other causes. E.g., being immunocompromised.
  - RD for immunocompromised status > 0; PAF>0.
- $PAF_{TB-exposure} + PAF_{immunocompromised} > 100$

This concept can be tied to Rothman's causal pies model of causality Epidemiology by design by Daniel Westreich

A measure related to PAFs

The **population attributable risk difference** is the difference between the risk of the outcome in the observed population and the risk of the outcome **if all exposure were removed**, that is:

$$P(Y) - P(Y^{x=0})$$

-The potential outcomes notation is *deliberate* here.

Epidemiology by design by Daniel Westreich

Remember our anaphylaxis by sex example?

Assuming a cohort design:

```
pcidat3<-c(687, 1144002, 462, 658943)

pciOR3<- epi.2by2(pcidat3, method = "cohort.count")
kbl(pciOR3$massoc.summary, digits = 2)
```

var	est	lower	upper
Inc risk ratio	0.86	0.76	0.96
Inc odds ratio	0.86	0.76	0.96
Attrib inc risk *	-0.01	-0.02	0.00
Attrib fraction in exposed (%)	-16.74	-31.35	-3.76
Attrib inc risk in population *	-0.01	-0.01	0.00
Attrib fraction in population (%)	-10.01	-18.04	-2.52

Attributable Risk ≅ Risk Difference

Number Needed to Treat

The number needed to treat (NNT) is the number of individuals who we would need to treat in order to **prevent** one bad outcome.

The NNT is calculated as $|RD|^{-1}$

- The inverse of the absolute value of the risk difference.
  - For a harmful exposure, we **keep the absolute value**, but describe the measure as a ***number needed to harm***.

Epidemiology by design by Daniel Westreich

NNT

If the 5-year risk of death among treated is 10%, and among untreated is 20%,

- then how many people do you need to treat to prevent one death over five years?
- $|10\% - 20\%|^{-1}$ = $|-10\%|^{-1} = 0.10^{-1} = 10.$
  - We must treat 10 people to prevent one death over five years.

Notation: $NNT = 1/|(P(Y|X = 1) - P(Y|X = 0))|$

What's the null value?

Epidemiology by design by Daniel Westreich

NNT

Often by convention, round NNT up to nearest integer - Can't treat half a person

- Conservative approach; doesn't always make sense.

- NNT implies causality more strongly than risk difference.
- Risk differences can be viewed as descriptive (although you should be specific and cautious about that)
- Just the difference in observed risks between two groups: not necessarily due to group identification.
- In contrast, NNTs explicitly discuss a treatment having a result: thus are expressing a causal effect.
- Explaining an NNT can become tricky outside of a trial setting because exposure (smoking) isn't always the same as treatment (cognitive behavioral therapy for smoking cessation).

Epidemiology by design by Daniel Westreich

QUESTIONS?

COMMENTS?

RECOMMENDATIONS?

5 emergency steps
for treating anaphylaxis

Food Allergy Canada

If an **anaphylactic reaction** happens, follow these steps:

Image courtesy of Kids' CAP

Give an **epinephrine auto-injector** (e.g., EpiPen®) right away. Follow the instructions on the device.

Call 9-1-1 or your local EMS immediately and tell them someone is having an anaphylactic reaction.

Image courtesy of Kids' CAP

Use a **second auto-injector** as early as 5 minutes after giving the first dose if there is no improvement in symptoms.

Go to the nearest **hospital** right away (ideally by ambulance), even if symptoms are mild or have stopped. The reaction could get worse or come back.

Call **emergency contact** person (e.g., parent, guardian, spouse).

The allergic reaction is the reason for going to the hospital, not because epinephrine has been used.

Incidence Odds Ratio

- For example: Define a population where 10% of people have $r_{1i} = 0.60$ and
- $r_{0i} = 0.20$ and 90% of the people have $r_{1i} = 0.035$ and $r_{0i} = 0.006$
- Here, the individual ORs = 6.0 for every individual
- the average of the individual ORs = 6.0
- Also, the ratio of the average odds equals 6.0 as well

But, the incidence odds ratio is equal to 3.9

Incidence Odds Ratio - Calculations

Rudimentary calculations

```
#10% pop
 0.6/(1-.6) #= 1.5 #r1
 0.2/(1-.2) #= 0.25 #r0

#OR in 10%
1.5/0.25 #= 6

#OR 90%
 0.035 / (1-0.035) # = 0.03626943 #r1
 0.006/(1- 0.006) # = 0.006036217 #r0

#OR in 90%
 0.03626943/0.006036217 #= 6.008636

# OR average
 1.5+0.03626943 #= 1.536269
 1.536269/2 #=   0.7681345
 0.25+0.006036217 #= 0.2560362
 0.2560362/2 #= 0.1280181

 0.7681345/0.1280181 #= 6.000202
```

```
## IOR
(0.6*0.1) + (0.035 *0.9) #= 0.0915
(0.2*0.1) + (0.006*0.9) #= 0.0254
#OR1
0.0915/(1- 0.0915) #=0.1007155
#OR0
0.0254/(1-0.0254) #=0.02606197
#IOR
0.1007155/0.02606197 #3.864462
```