Measures of Occurrence

Mabel Carabali

EBOH, McGill University

Updated: 2025-09-07

Objectives

- To review core concepts of measures of occurrence
- To clarify potential misconceptions about measures of occurrence
- To identify opportunities to improve the reporting of measures in epidemiology

Chapters 4 and 5: Rothman, K. J., Lash, T. L., VanderWeele, T. J., & Haneuse, S. (2021). Modern epidemiology (Fourth edition). Wolters Kluwer. Available here

Expected competencies

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

Montreal

A teen died from an allergic reaction. Urgences-santé now owes his family $450K

Justice Jeffrey Edwards ruled that ambulance technicians failed to follow protocol

Rachel Watts · CBC News · Posted: Sep 05, 2025 4:41 PM EDT | Last Updated: September 5

https://www.cbc.ca/news/canada/montreal/allergic-reaction-death-quebec-judge-1.7626503

Some facts...

Here are some staggering facts:

Pre-hospital use of epinephrine is low:

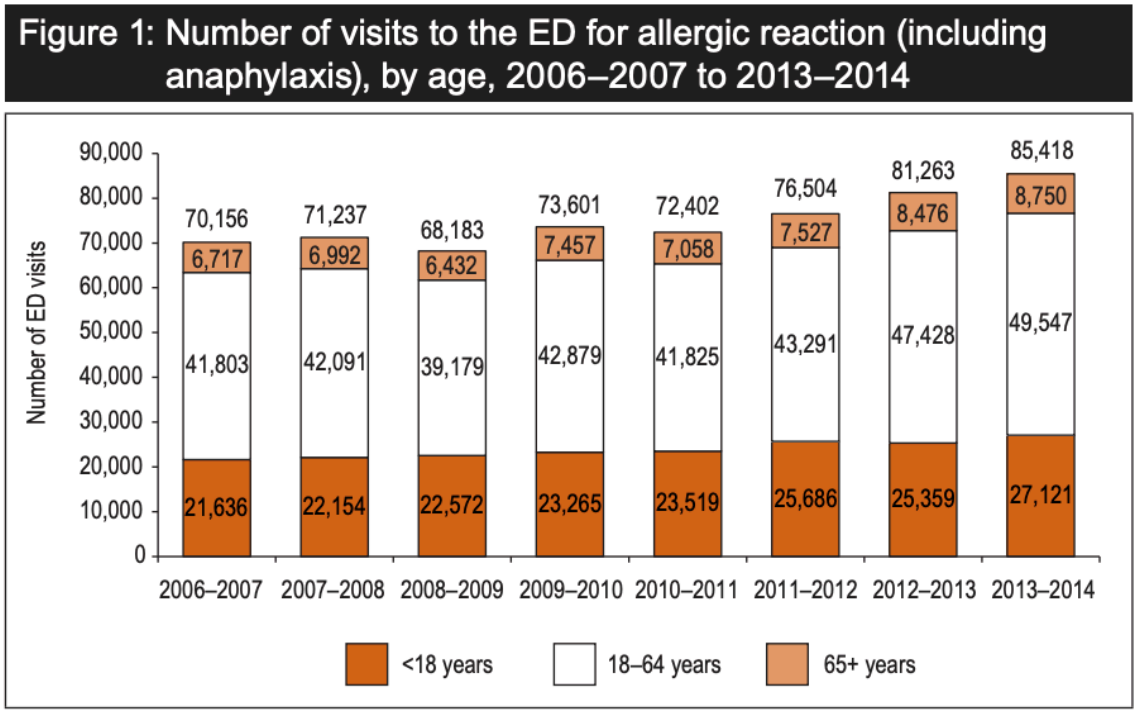
- EMS usage of epinephrine to treat anaphylaxis in Canada is only 36%
- Only 21% of children and 7% of adults globally use their auto-injector prior to going to the hospital
- Prompt use of epinephrine to treat anaphylaxis improves health outcomes, yet only 1-in-5 children and less than 1-in-10 adults use it before they go to the hospital

Anaphylaxis management issues:

- 25% of first-time reactions occur at school/daycares
- Children who did not receive prompt administration of epinephrine were more than 2x more likely to be admitted to the hospital

https://foodallergycanada.ca/campaign/food-allergy-and-the-holidays-2024/knowing-how-and-when-to-treat-anaphylaxis/

"Each year, 1% of all emergency department (ED) visits are for an allergic reaction (including anaphylaxis). 8% of these visits are for anaphylaxis specifically."



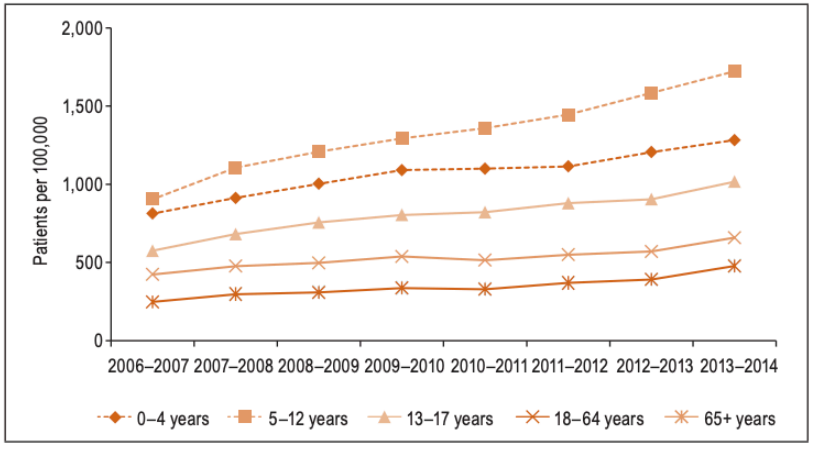
Note: Includes Ontario and Alberta. Sources: National Ambulatory Care Reporting System, 2006–2007 to 2013–2014, Canadian Institute for Health Information; Alberta Ambulatory Care Reporting System, 2006–2007 to 2009–2010, Alberta Health Services. CIHI Report

No Updated Data???

> "*The **prevalence** of self-reported food allergy increased from 7.1% to 9.3% between 2010 and 2016, whereas food allergy based on history or physician diagnosis remained stable (5.9% vs 6.1%). This increase in self-reported allergy is likely attributable to increasing awareness.*"

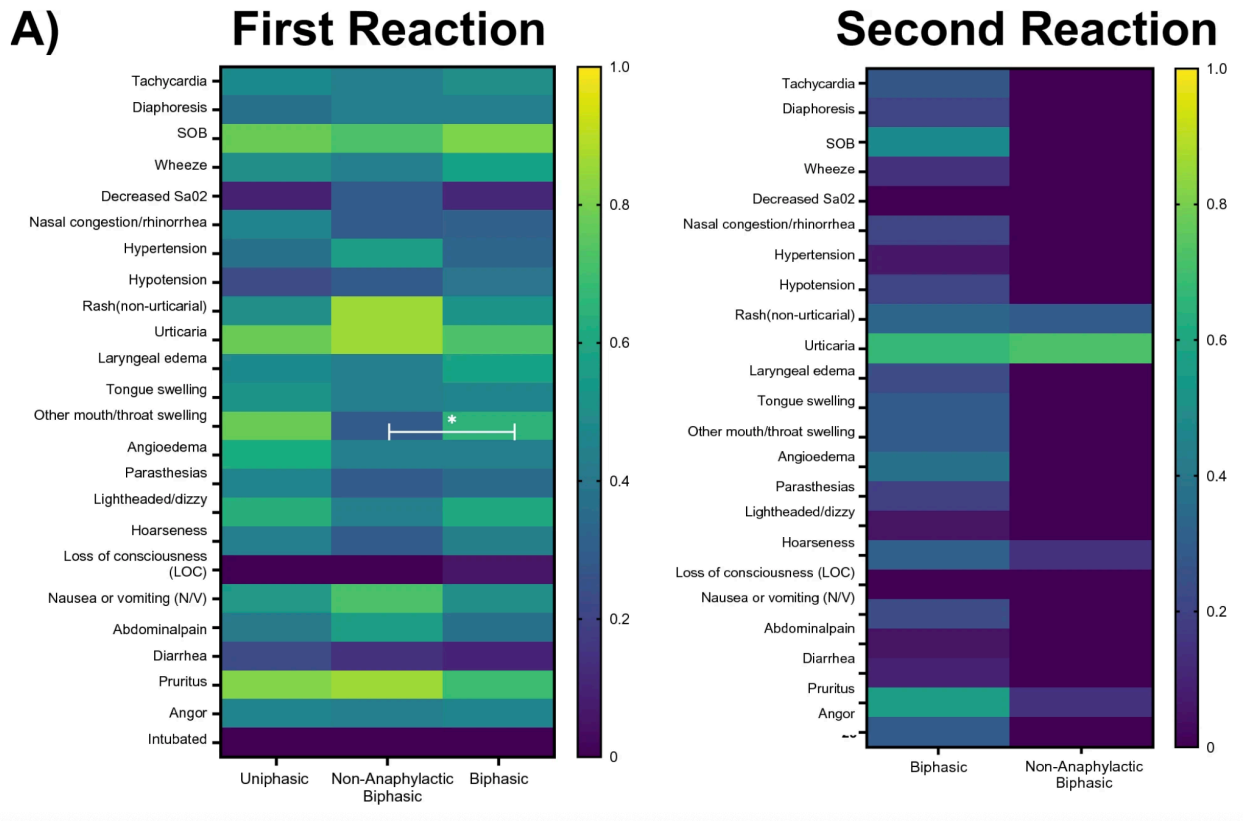
Clarke, Ann E. et al.(2020) Temporal trends in prevalence of food allergy in Canada. The Journal of Allergy and Clinical Immunology: In Practice, Volume 8, Issue 4, 1428 - 1430.e5

Figure 3: Rates of dispensed prescription auto-injectors per 100,000, by age, 2006–2007 to 2013–2014



*"Since 2006–2007, there has been a 64% increase in the **rate** of individuals dispensed an epinephrine auto-injector."* CIHI Report

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



Symptoms profiles. A Heat map (left) proportional occurrence of specific symptoms across the first anaphylactic response of uniphasic, non-anaphylactic biphasic reaction responders. (right) proportional occurrence of specific symptoms across the second anaphylactic response of non-anaphylactic biphasic and biphasic responders

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

> Methods: *Patients were contacted sometime after ED discharge to obtain consent and confirm symptoms and timing of the reaction*
>
> Data Analysis: *Statistical analyses were completed using GraphPad Prism 9.*
>
> Results: *Of 138 anaphylactic events identified, 15.94% were biphasic reactions, 79.0% were uniphasic, and 5.07% were classified alternatively as a non-anaphylactic biphasic reaction.*
>
> Conclusion: *The **incidence** of biphasic reactions in this cohort was 15.94% and the average second-phase onset was 19.0 h. In biphasic reactivity, it appears that the symptom profile second reaction is less severe compared to the first reaction.*

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

Reproduced Table 1

Comparator	Uniphasic (n=109)	Non-anaphylactic biphasic (n=7)	Biphasic (n=22)	P value
Age, median y	31.00	20.00	31.50	0.8565
Pediatric cases (<13 y), No. (%)	9 (8.3%)	-	1 (4.5%)	-
Females, No. (%)	68 (62.4%)	5 (71.4%)	14 (63.6%)	0.8900
History of anaphylaxis, No, (%)	58 (53.2%)	3 (42.9%)	14 (63.6%)	0.5528
History of Asthma, No (%)	43 (39.4%)	3 (42.9%)	7 (31.8%)	0.7483
Time to onset of symptoms for first reaction, minutes	N=90	N=6	N=17	0.9793
Median	10.00	10.00	15.00	
Mean	31.21	21.17	35.82	
Time to resolution of symptoms for first reaction, hours	N=97	N=6	N=19	0.0677
Median	3.750	5.500	2.500	
Mean	6.084	6.250	9.487	

Measures of Frequency

Occurrence of an outcome, generally/broadly presented as: **Number of individuals with the outcome (health or otherwise)**

Measures of Frequency (Epidemiology)

Distribution of disease [1] frequency in human populations

- Description, Surveillance
- Etiology, Causation

[1] *Any health outcome*

Measures of Frequency

What's the information we need ?

- Outcome of interest (Quantity)
- Time
- Space

POPULATION AT RISK

Measures of Frequency

- Number of individuals with the outcome **[Numerator]**
- Number of individuals at risk for that outcome **[Denominator]**

Next Question(s)

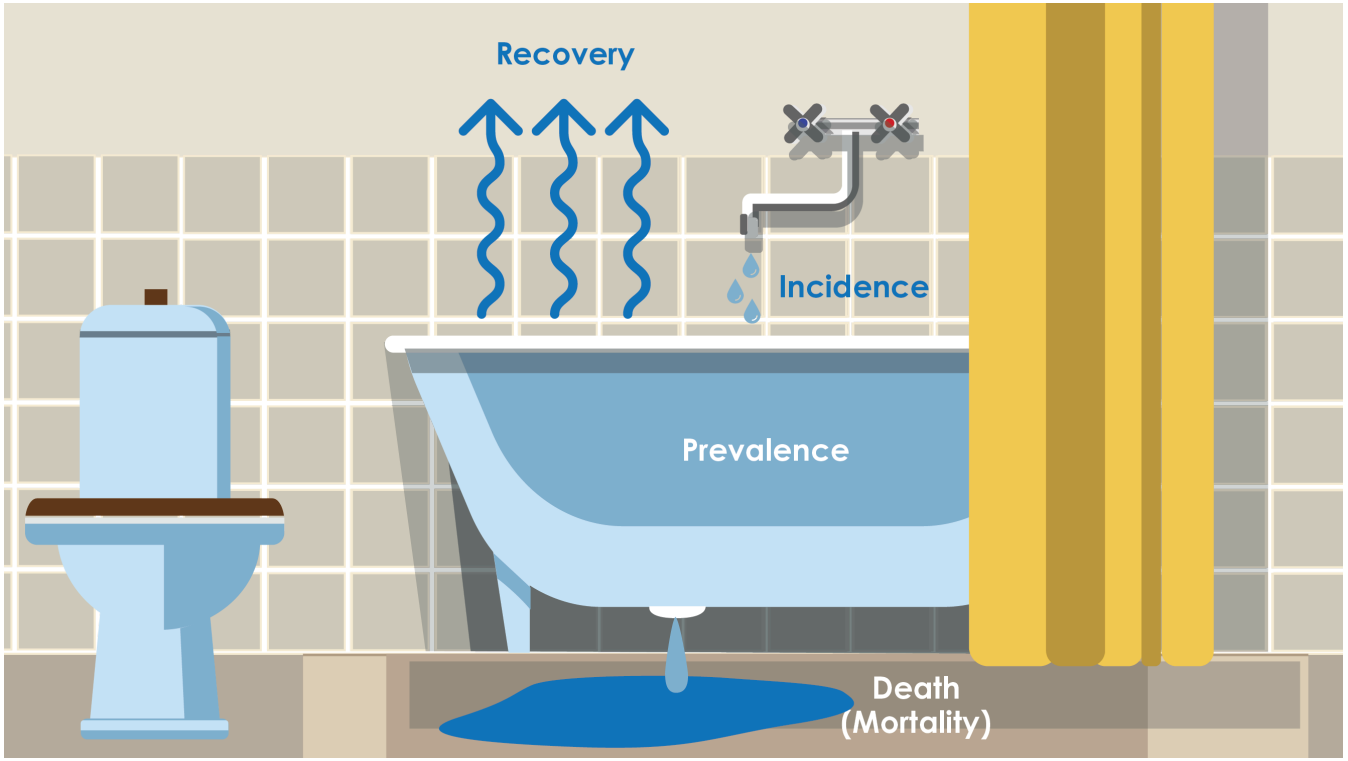
- Newly Developed Outcomes
- Existent Outcomes

Measures of Frequency

- Incidence (***New events***)
    - Incidence Proportion
    - Incidence Times
    - Incidence Rates
- Prevalence (***Existing Events***)

Incidence vs Prevalence

New/existing Outcomes the same?



https://www.technologynetworks.com/immunology/articles/incidence-vs-prevalence-329073

What about the *Population* ? 🤷

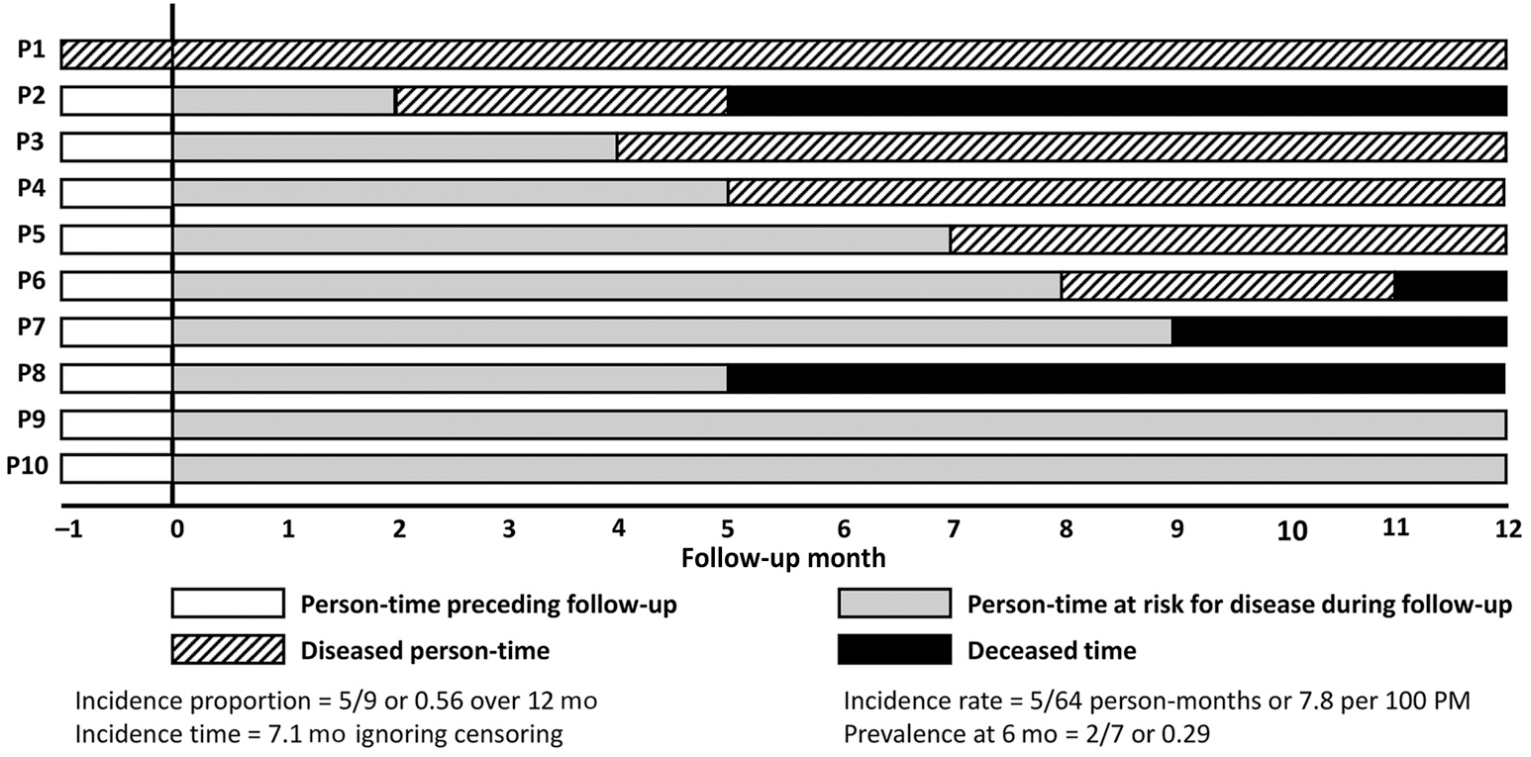


Illustration of four basic measures of disease frequency using a population of 10 persons (P1-P10) with person-time depicted 1 month before, and 12 months after, the start of follow-up for the occurrence of a disease. Lash, Timothy, L. et al. Modern Epidemiology. Available from: Wolters Kluwer, (4th Edition). Wolters Kluwer Health, 2020.Chapter 4, page 54

Measures of Frequency

- Incidence (*New events*)
  - Incidence Proportion (*Cumulative Incidence*, $Risk$[1], Proportion_)
  - Incidence Times (*Event time, failure time, occurrence time*)
  - Incidence Rates (*Incidence density*, $Hazard$[2] */ Hazard rate*)
- Prevalence (*Existing Events*)

[1] Average Risk
[2] Instantaneous Hazard

A note on ***Statistical Lexicon:*** https://statmodeling.stat.columbia.edu/2009/05/24/handy_statistic/

Incidence Proportion

I = *Incidence proportion*

a.k.a Cumulative incidence

A = *Number of New Outcomes*

N = *Initial Size of the Population*

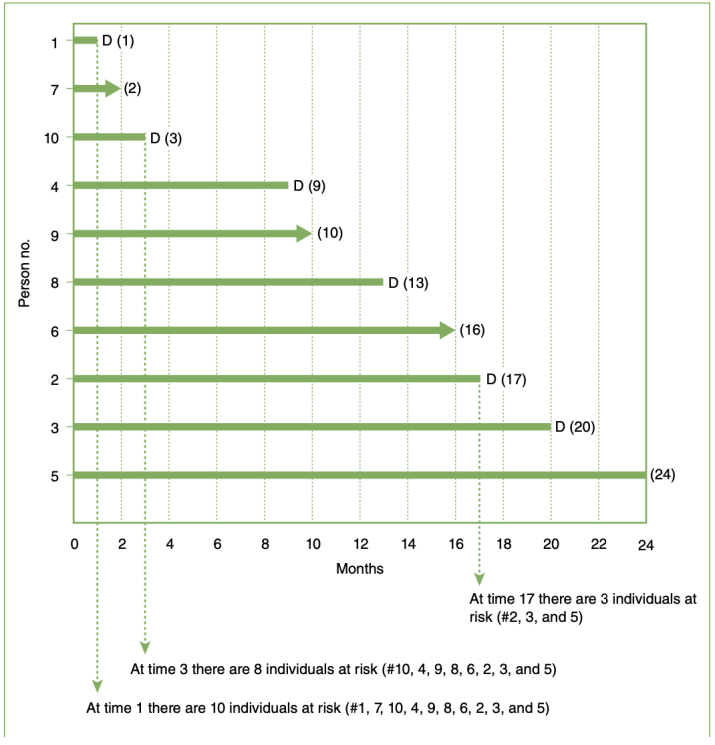
$$I = \left(\frac{A}{N}\right)$$

- Range: 0 to 1
- Dimensionless
- **Must state the time period for which it applies**

From: Modern Epidemiology 4th Ed. Ch 4 & Epidemiology Beyond the Basics 3rd Ed. Ch2

Incidence Proportion

Example:



Individuals sorted according to follow-up time from shortest to longest. D = death; Arrow (C)= censored observation; ( ), duration of follow-up in months (all assumed to be exact whole numbers).

Incidence Proportion

$I = \left(\frac{A}{N}\right)$ **I = 6/10 = 0.6**

Assumes a ***stable/static*** population:

Drawback: *"over any appreciable time interval, it is usually* ***technically impossible to measure risk.*** *The reason is a practical one: if a population is followed over a period of time, some people in the population will die from causes other than the outcome under study."* -Rothman, pg. 26

- **Competing risks:** Occurs when other causes (non related to the outcome of interest) prevent the participants to develop the outcome and affect follow-up time *E.g., the patient died in a car crash when the outcome under study is MI or cancer.*

- **Loss to follow-up:** Lack complete information for the intended period of follow-up time *E.g., the study participant withdraw from an RCT after 6 months in a study expected to last 12 months.*

So, what if...

Hypothetical experiment: randomized 100 people to each exposure condition

- Exposure: smoking vs. non-smoking (or allergens)
- Outcome: death or anaphylaxis

Do you expect to see a difference in the number of deaths by exposure condition?

- Over the next week? 3 minutes
  - Over the next 25 years? 3 hours
  - Over the next 100 years? 3 days

Capture the etiologically relevant time period

So, time matters...

Incidence Times (*Event time, failure time, occurrence time*)

Time span from time zero to time of the event

- Will be different when using different time scales

Incidence Times

When do we start measuring incidence times (i.e., time zero)?

- From birth (when age at event is used)
- From study entry (in an RCT, for example)
- A specific date (e.g., on 9/11 in a study of PTSD of NY firefighters)

- Example: A 80 year old individual has a stroke.
    - What is the incident time?

It depends !

- On the age scale?

- Now, what would it be if was enrolled in an RCT at age 65 ?
    - Using years since study entry

- What is the incident time if the stroke does not occur by age 80, but the RCT ends when the participant is 80 years old? i.e., Censoring
    - Using years since study entry

Censoring

Incomplete follow-up time.

An **incidence time** is censored if it is possible that the outcome of interest would have taken place after the point in time that censoring occurred.

Common censoring reasons:

- End of study follow-up
- Loss to follow-up/Withdraws
- Competing risk*

Censoring

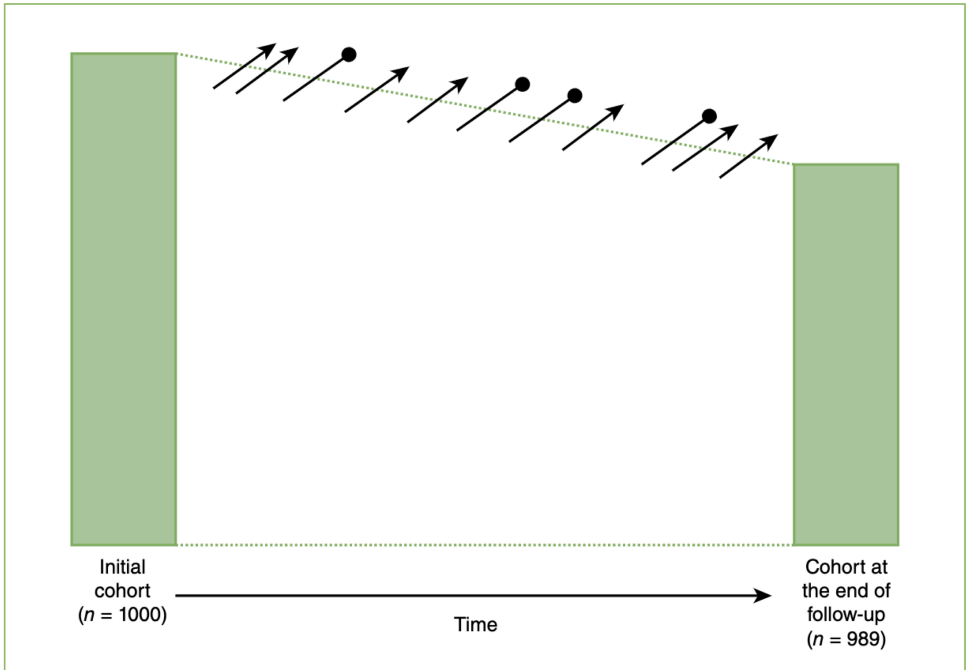


FIGURE 1-13 Diagram of a hypothetical cohort of 1000 subjects. During the follow-up, four disease events (line segments ending in dots) and seven losses to follow-up (arrows) occur so that the number of subjects under observation at the end of the follow-up is 989.

Szklo M, Nieto FJ. Epidemiology : Beyond the Basics. Fourth ed.pg 21

Incidence Times

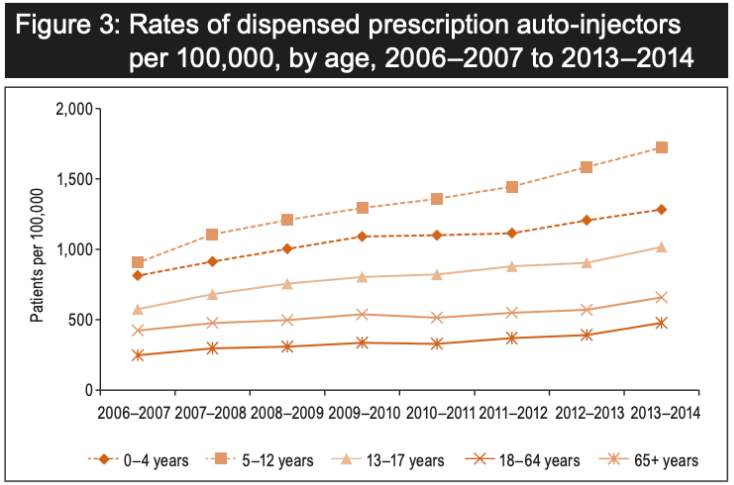
Descriptive (when there is no censoring)

- Mean incidence time (life expectancy)
- Median incidence time

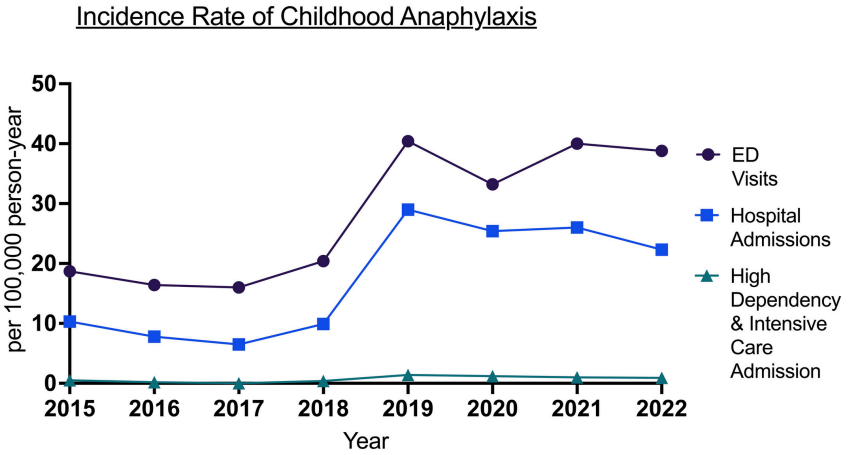
What is the difference between life expectancy and the so-called “average age at death” ?

- Which one is calculated at a given time point?
- Which one is more relevant when considering etiology?

Rates vs *Incidence Rates*



Anaphylaxis and Allergy in the Emergency Department (Canada)



Trends in Childhood Anaphylaxis in Singapore: 2015–2022

Incidence Rates

- Very important concept in epidemiology: **Person-time at risk.**
- Addresses censoring (time issues)
    - e.g., the technical problem with measuring the incidence proportion under competing risks.

"The total person-time at risk merely represents the total of all time during which disease onsets could occur and would be considered events occurring in the population of interest."

Drawback of rate: **harder to interpret, less intuitive**

Counting Person-Time

In a study of uterus cancer (O)

- Should a participant with a $hysterectomy$ (✕) at the start of your study contribute to the denominator?

  ✕|----------|

- What if the participant has a hysterectomy during follow-up time?

  |------- ✕ --------|

  |----------| ?? ✕ ??

  - Only outcomes that occur to persons contributing to the denominator at the time that "disease" onset occurs should be counted in the numerator.
  - Only people that can possibly have an outcome should contribute person-time to the denominator

Counting Person-Time

- Consider patients that need heart transplants and a study of whether the transplant will impact mortality
- In a study of this question, the author contrasted the incidence rate in those that received the transplant to those that did not receive the transplant but were waiting for one. All the person time from the group that received the transplant was classified as **exposed**, including the **person time before the transplant was received**.
- This artificially makes the mortality rate in the exposed group lower than it should be because before they received the transplant they could not have died. **If they did, then they would have been classified as unexposed.**

Definition of **"Immortal time"**: span of time when the outcome could not have occurred

- Usually occurs with passing of time before a subject initiates exposure -To be classified as exposed, the subject MUST remain event-free until the start of exposure

Immortal Time Bias (Suissa 2007)

How do we account for the time?

Incidence Rates

IR = *Incidence Rate*

A = *Number of New Outcomes*

$\overline{N}$ = *Average Population over the study period*

$\Delta\ t$ = *Length of the risk period*

$\overline{N} \cdot \Delta\ t$ = *Total Person Time at Risk*

$$IR = \left(\frac{A}{\overline{N} \cdot \Delta\ t} \right)$$

Incidence Rates

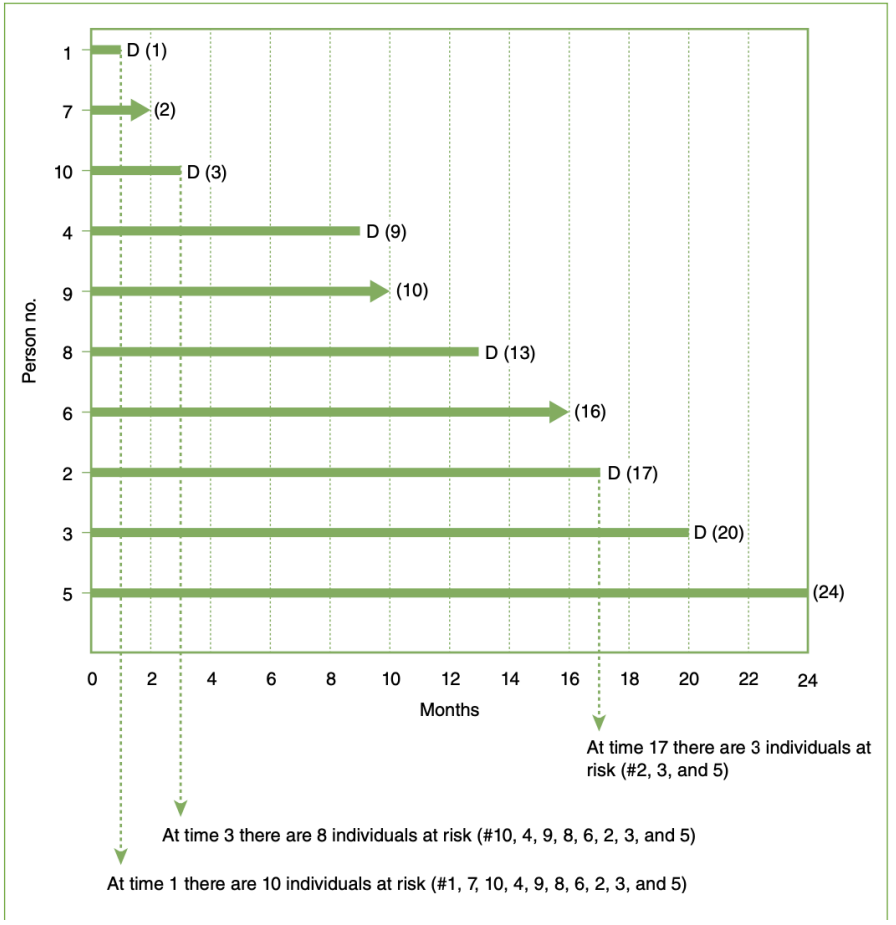
As a time weighted average of individual rates

A^* = $\sum_{people}(Time\ spent\ in\ population)(Individual\ Rate)$

$$IR = \left(\frac{\sum_{people}(Time\ spent\ in\ population)(Individual\ Rate)}{\sum_{people}(time\ spent\ in\ population)} \right)$$

$$IR = \left(\frac{A^*}{\sum_{people}(Time\ spent\ in\ population)} \right)$$

Hands on!



Individuals sorted according to follow-up time from shortest to longest. D = death; Arrow (C)= censored observation; ( ), duration of follow-up in months (all assumed to be exact whole numbers).

Incidence *Rate* vs Incidence *Density*

Incidence Rate Based on Aggregate Data

$\overline{N}$ = *Average Population over the study period (yrs)*

Incidence Density Based on Individual Data

ID = Incidence Density

$$ID = \left(\frac{A}{(Total\ person\ time\ )} \right)$$

- Total Person time = Total follow-up time period = $\overline{N} \cdot \Delta\ t$ = *Total Person Time at Risk*

Incidence Rate Based on Aggregate Data

$\overline{N}$ = Average Population over the study period (yrs)

```
d = 6 # number of deaths
n1 = 10 - 0.5*(6+3) # half of the events and losses minus initial population
n2 = (10 +1) /2 # 10 people at time zero + 1 alive at the end (2yrs)
cbind(n1, n2)
```

```
##       n1  n2
## [1,] 5.5 5.5
```

$IR_{\overline{N}}$ = d (n=6) / ($\overline{N}$ = 5.5); IR = (1.09) per 2 person-years

- Also, 1.09 ÷ 2 years = 0.545 per person-year

- or $IR_{\overline{N}}$ **= 54.55 per 100 person-years**.

Incidence Density Based on Individual Data

Person Time

```
d = 6 #number of deaths
#deaths, time contributed
Dt= 1 + 3 + 9 + 13 + 17 + 20
Dt
```

```
## [1] 63
```

```
Ct= 2 + 10 + 16 #censoring,time contributed
Ct
```

```
## [1] 28
```

```
Et = 24 # complete follow-up, time contributed
PT <- Dt + Ct +Et #time contributed
PT
```

```
## [1] 115
```

```
IR <- 6/PT
IR
```

```
## [1] 0.05217391
```

The IR_{pt} = Events (6)/ Person-Time (115);

IR =0.052, per person-months

Incidence *Density*

Recall: Incidence Density Based on Individual Data

ID = Incidence Density = IR based on ***Individual Person-time*** $\overline{N} \cdot \Delta t$

Person time = 115 for 24 months (2-years), what's for a year?

```
PT
```

```
## [1] 115
```

```
TFU <- PT/12 #months
TFU
```

```
## [1] 9.583333
```

```
ID <- d/TFU
ID
```

```
## [1] 0.626087
```

Incidence Density

$\overline{N} \cdot \Delta t$ = ***Total Person Time at Risk***

ID = Incidence Density = IR based on ***Individual Person-time***

ID = Events (n= 6) / Total Follow-up person time (n= 9.5833333)

ID **= 0.63 per person-year, or 62.61 per 100 person-years**

ID = D (n=6) / TFU (9.583 × 12 months) = 0.052, or 5.2 per 100 person-months

ID= 6/115 = 0.052= or 5.2 per 100 person-months

Incidence *Rate* or Incidence *Density*

Assumptions:

- Independence between censoring and survival and of lack of secular trends
- The risk of the event remains approximately constant over time during the interval of interest *"The estimated rate should apply equally to any point in time within the interval"*
- *n* people followed during *t* units of time are equivalent to *t* people observed during *n* units of time;

  BUT 1 smoker followed for 30 years **is certainly not the same** as that of 30 smokers followed for 1 year

Incidence *Rate* or Incidence *Density*

Person no.	Total follow-up (in months)	Contribution to the total number of person-years by participants in:		
		1st Year of follow-up	2nd Year of follow-up	Total follow-up period
1	1	1/12 = 0.083	0	0.083
2	17	12/12 = 1.000	5/12 = 0.417	1.417
3	20	12/12 = 1.000	8/12 = 0.667	1.667
4	9	9/12 = 0.750	0	0.750
5	24	12/12 = 1.000	12/12 = 1.000	2.000
6	16	12/12 = 1.000	4/12 = 0.333	1.333
7	2	2/12 = 0.167	0	0.167
8	13	12/12 = 1.000	1/12 = 0.083	1.083
9	10	10/12 = 0.833	0	0.833
10	3	3/12 = 0.250	0	0.250
Total	115 months	7.083 years	2.500 years	9.583 years

Szklo M, Nieto FJ. Epidemiology : Beyond the Basics. Fourth ed.pg 68

Incidence *Rate* or Incidence *Density*

Option: Divide the follow-up into smaller intervals and calculate incidence densities for each interval

- First follow-up year: 3 ÷ 7.083 = 42.4 per 100 person-years (or 3 ÷ 85 = **3.5** per 100 person-months)

- Second follow-up year: 3 ÷ 2.500 = 120 per 100 person-years (or 3 ÷ 30 = **10** per 100 person-months)

Given the differences, *not advisable* to estimate an incidence density for the overall 2-year period.

Incidence *Rate* or Incidence *Density*

- It is of practical interest that when withdrawals (and additions in an open population or dynamic cohort) and events occur uniformly, rate (based on grouped data) and density (based on individual data) are virtually the same.

- The following equation demonstrates the equivalence between the rate per average population and the density (per person-time), when the former is averaged with regard to the corresponding time unit (e.g., yearly average).

$$Rate(IR_{\overline{N}}) = \left(\frac{\frac{No.events(x)}{average\ population(n)}}{time(t)} \right) = \left(\frac{x}{n \times t} \right) = Density(ID)$$

But it is NOT always the same!

Incidence *Rate* or Incidence *Density*

$IR_{\overline{N}}$ = 54.55 per 100 person-years $\neq ID$ = 62.61 per 100 person-years

$IR_{\overline{N}}$ = Events (n= 6)/ Average Population over time (n= 5.5)

$IR_{\overline{N}}$ = 54.55 per 100 person-years

ID = Events (n= 6) / Total Follow-up period (Person-Time) (n= 9.583)

ID = 62.61 per 100 person-years

"When the sample size is large and provided that the time interval is reasonably short, the assumption of uniformity of events/losses is likely to be met" (Szklo M, Nieto FJ. Epidemiology : Beyond the Basics. Fourth ed.

Incidence *Rate* or Incidence *Density*

Range: between 0 and infinity

- Must report the time units used! Person-years, person-months, etc because the incidence rate can change drastically based on the unit – don't confuse yourself or your reader!

Example 1/100PYs – how many people were observed for how long?

Incidence *Rate* or Incidence *Density*

Caveat: Assumed to be constant for the time window in which they are measured.

- 10 persons followed for 100 years vs. 1000 persons followed for 1 year – These will usually show different incidence rates of events

- Clearly define the time windows of observation.

If it isn't likely to show a constant incidence rate then divide it into finer strata and calculate separate rates for each strata.

Incidence Proportion, Incidence Rate and Incidence Density

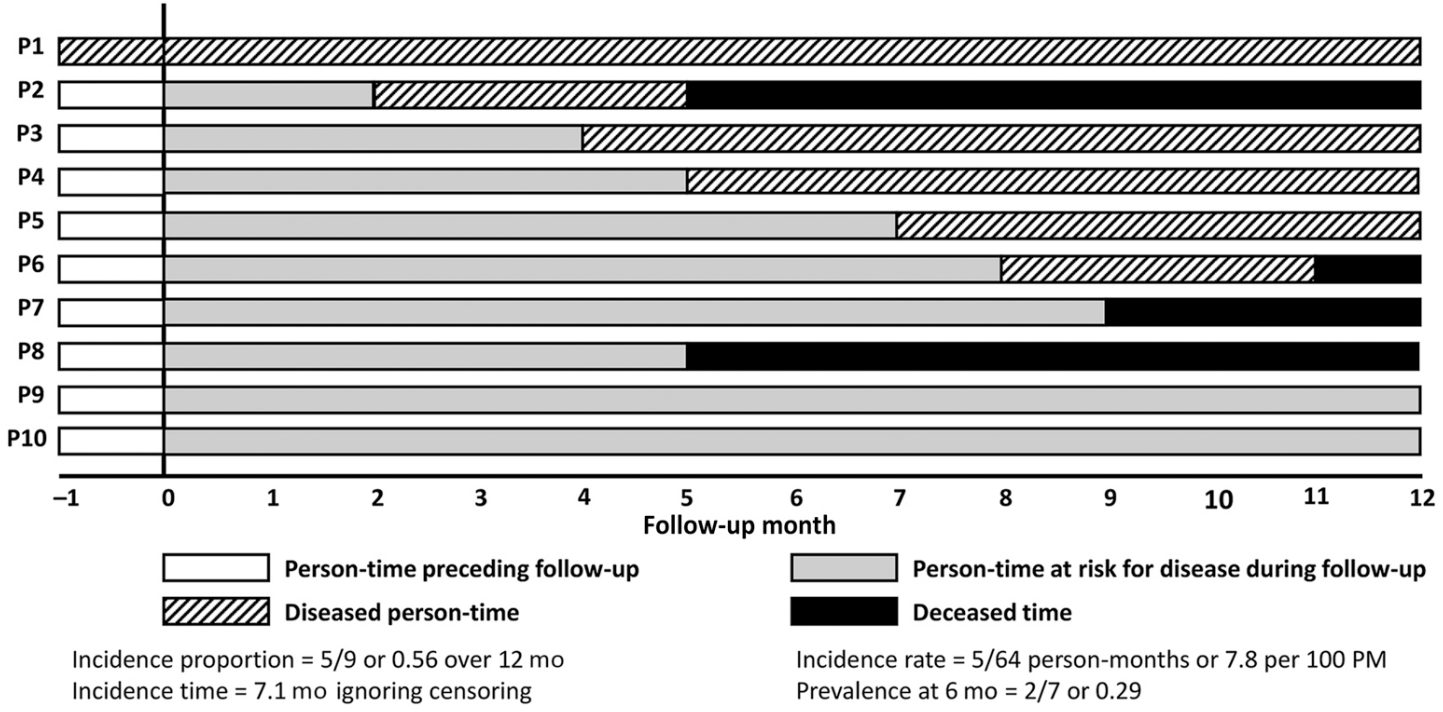


Illustration of four basic measures of disease frequency using a population of 10 persons (P1-P10) with person-time depicted 1 month before, and 12 months after, the start of follow-up for the occurrence of a disease. Lash, Timothy, L. et al. Modern Epidemiology. Available from: Wolters Kluwer, (4th Edition). Wolters Kluwer Health, 2020.Chapter 4, page 54

Incidence proportion (*risk*) and Incidence rate

	Incidence Proportion	Incidence Rate
Smallest value	0	0
Greatest value	1	Infinity
Units (dimensionality)	None	1/time
Interpretation	Probability or average risk	Inverse of waiting time (under steady state conditions)

(Basic) Simulated examples

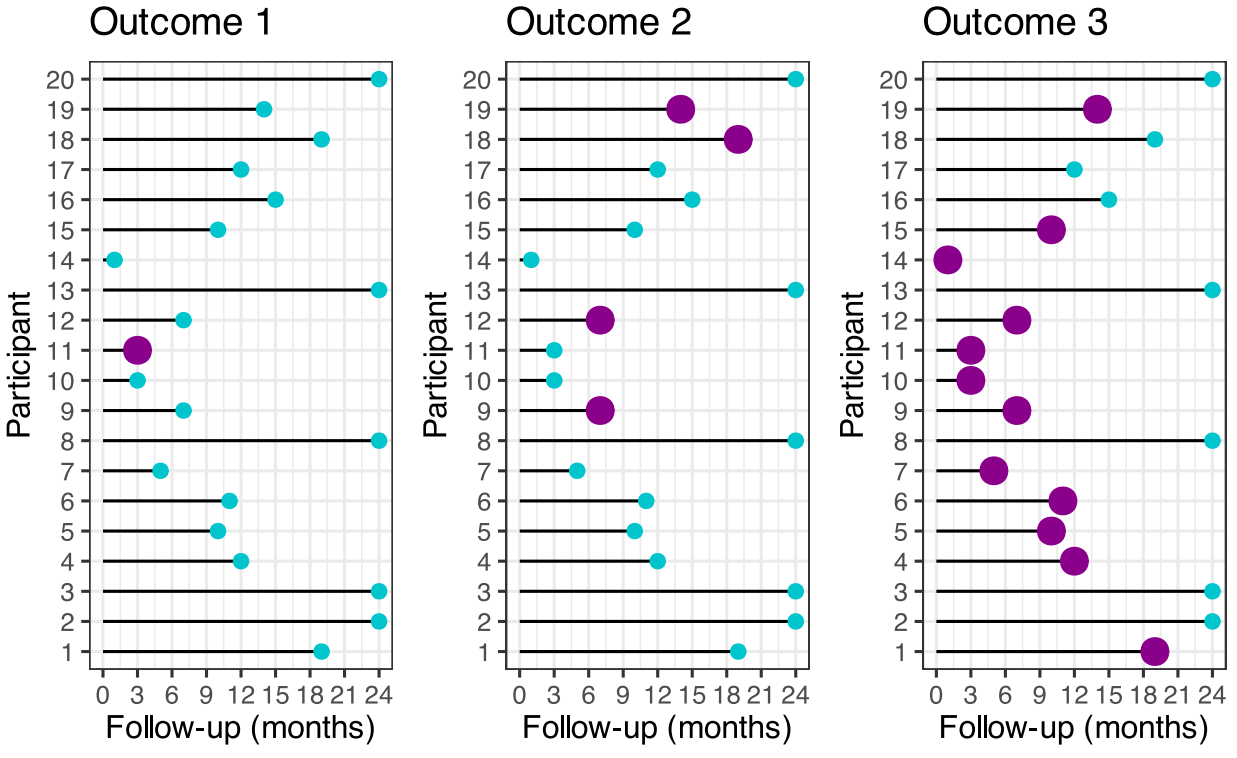
```
set.seed(7042025)
O1<- rbinom(20, 1, 0.10)
O2<- rbinom(20, 1, 0.45)
O3<- rbinom(20, 1, 0.75)
t <- sample(1:24, 20, replace = TRUE)
t[t>=20]<-24


dat<-cbind.data.frame(ID=1:20,O1, O2, O3, t, nb=1)
dat <- dat %>%
  mutate(O1=
           ifelse(O1==1 & dat$t==24, 0, O1),
           O2= ifelse(O2==1 & dat$t==24, 0, O2),
           O3= ifelse(O3==1 & dat$t==24, 0, O3))
```

Example Table

ID	O1	O2	O3	t	nb
1	0	0	1	19	1
2	0	0	0	24	1
3	0	0	0	24	1
4	0	0	1	12	1
5	0	0	1	10	1
6	0	0	1	11	1
7	0	0	1	5	1
8	0	0	0	24	1
9	0	1	1	7	1
10	0	0	1	3	1
11	1	0	1	3	1
12	0	1	1	7	1
13	0	0	0	24	1
14	0	0	1	1	1
15	0	0	1	10	1

Different Scenarios



- **Magenta** Dead
- Turquoise: Alive/Censored

What do we have now?

Person-months = 268 ; **Person-years** = 22.333; $\overline{N}$ = 12.5

Outcome 1

- nb outcomes = 1
- IP_{O1} = 0.05
- IR_{O1} = 0.04 per person-years
- ID_{O1} = 0.04 per person-years

Outcome 2

- nb outcomes = 4
- IP_{O2} = 0.2
- IR_{O2} = 0.16 per person-years
- ID_{O2} = 0.18 per person-years

Outcome 3

- nb outcomes= 12
- IP_{O3} = 0.6
- IR_{O3} = 0.48 per person-years
- ID_{O3} = 0.54 per person-years

Relationship between Incidence Rate and Incidence Proportion

If population is **closed**, and the time interval is short enough that the size of the population only declines slightly over the interval.

$$Incidence\ Proportion\ \cong Rate \times \Delta\, t$$

But! this approximation doesn't hold for diseases with "high" risks because they would involve more than a few cases occurring and the approximations that the formula relies on would not hold.

A note on Populations

Open

- May gain members over time
- Lose members to death or some other means (loss to follow-up, emigration)

Closed

- Adds no new members over time
- Lose members only to death

Type of population depends on the time axes chosen;

- A population can be closed using one axis and open using another

Open vs. Closed Populations

- Open vs. closed is an important distinction and can have important implications especially when designing a study.

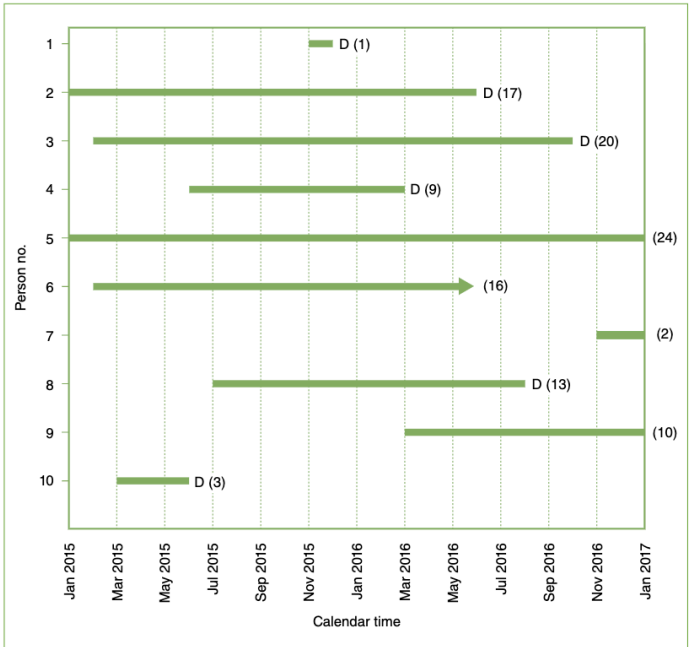


FIGURE 2-1 Hypothetical cohort of 10 persons followed for up to 24 months from January 2015 through December 2016. D, death; arrow, censored observation; ( ), duration of follow-up in months (all assumed to be exact whole numbers).

(Szklo M, Nieto FJ. Epidemiology : Beyond the Basics. Fourth ed.)

Population

- Temporal
- Dynamic

Examples: population of a town or country, epidemiological study of driving, cell phone use and accidents

Cohort

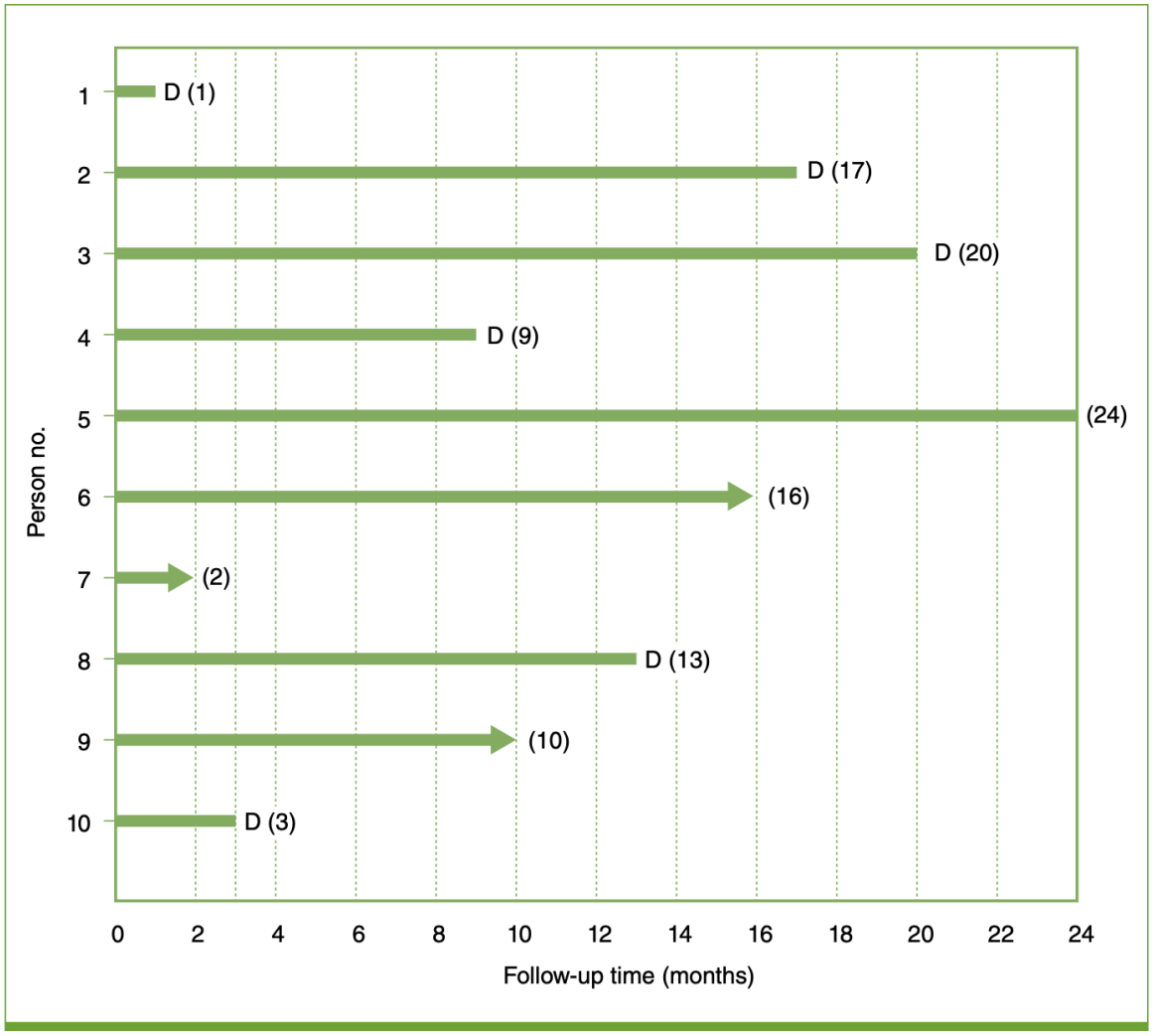
- Membership defined in a permanent fashion
- "Fixed"/Closed

Examples: birth cohorts, patients followed up from date of surgery

Relationship between Incidence Rate and Incidence Proportion

- The method of calculating risks over a time period with changing incidence rates is known as survival analysis.

- *"The cumulative probability of the event during a given interval lasting m units of time and beginning at time x, is the proportion of new events during that period of time in which the denominator is the initial population corrected for losses"*. (Szklo M, Nieto FJ. Epidemiology : Beyond the Basics. Fourth ed.)



(Szklo M, Nieto FJ. Epidemiology : Beyond the Basics. Fourth ed.)

Assumptions in the Estimation of Cumulative Incidence Based on Survival Analysis

- Uniformity of Events and Losses Within Each Interval (Classic Life Table).
- Events and losses are approximately uniform during each defined interval.
- If risk changes rapidly within a given interval, then calculating a cumulative risk over the interval is not very informative.
- The rationale underlying the method to correct for losses—that is, subtracting one-half of the losses from the denominator also depends on the assumption that losses occur uniformly.
- Independence of censoring AND Survival
- **No secular trends!**

The *odd* ODDS

Selected measure = $\left(\frac{Number\ of\ events}{Denominator}\right)$

Denominator Population at risk at the beginning or at the end or the person time.

The Odds are functions of probabilities (P) as follows

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

$P = 5; Odds = \left(\frac{0.05}{1-0.05}\right) = \left(\frac{0.05}{0.95}\right)$ = 0.0526316

- While probabilities are bounded, Odds are NOT

- If the probability is small/low the Odds are similar but... $> P > Odds$

A note on Odds

Example We observe four outcomes with four different probabilities each,

Outcome 1 = 1%, Outcome 2 = 5%, Outcome 3 = 35%, Outcome 4 =75%

```
Odds1 = 0.01/(1-0.01)
Odds2 = 0.05/(1-0.05)
Odds3 = 0.35/(1-0.35)
Odds4 = 0.75/(1-0.75)
```

$Odds_{O1}$ = **0.01** < $Odds_{O2}$ = **0.05** < $Odds_{O3}$ = **0.54** < $Odds_{O4}$ = **3**

More on the Oddity of the Odds later

Prevalence

- Range: [0,1], dimensionless
- Snapshot at a point (point prevalence) in time or during a specified period (period prevalence)
- Useful for measuring disease burden and for resource planning (how many people at this moment are suffering from outcome X ?)
- Impacted by both duration of the condition and the incidence rate

Note: Etiology of disease is interested in incidence (disease onset), so prevalence is a ***poor measure for studying etiology*** (but sometimes it is all we have!)

Recall this statement?

> "*The **prevalence** of self-reported food allergy increased from 7.1% to 9.3% between 2010 and 2016, whereas food allergy based on history or physician diagnosis remained stable (5.9% vs 6.1%). This increase in self-reported allergy is likely attributable to increasing awareness.*"
>
> *Methods: By **using 2006 Canadian Census data**, postal codes with high proportions of vulnerable populations were identified and households were randomly selected to participate in a telephone survey. Information on food allergies and demographics was collected. Prevalence estimates were weighted by using Census data to account for the targeted sampling.*
>
> *Results: Of 12,762 eligible households contacted, 5734 households completed the questionnaire (45% response rate).*

Clarke, Ann E. et al.(2020) Temporal trends in prevalence of food allergy in Canada. The Journal of Allergy and Clinical Immunology: In Practice, Volume 8, Issue 4, 1428 - 1430.e5

Clarke, Ann E. et al.(2020) Temporal trends in prevalence of food allergy in Canada. The Journal of Allergy and Clinical Immunology: In Practice, Volume 8, Issue 4, 1428 - 1430.e5

TABLE I. Weighted perceived prevalence estimates of individual foc

	All		
Allergen	S2S (n = 15,322)	SPAACE (n = 15,022*)	Difference†
Peanut	1.4 (1.2 to 1.7)	1.0 (0.8 to 1.3)	0.4 (−0.001 to 0.7)
Tree nut	1.8 (1.5 to 2.2)	1.3 (1.1 to 1.7)	0.5 (0.1 to 1.0)
Fish	0.8 (0.6 to 1.0)	0.7 (0.5 to 1.0)	0.1 (−0.2 to 0.4)
Shellfish	1.9 (1.6 to 2.2)	1.8 (1.5 to 2.1)	0.1 (−0.4 to 0.5)
Sesame	0.3 (0.2 to 0.5)	0.2 (0.1 to 0.3)	0.1 (−0.1 to 0.3)
Milk	2.6 (2.3 to 3.0)	0.7 (0.5 to 0.9)	1.9 (1.5 to 2.4)
Egg	0.9 (0.7 to 1.2)	0.6 (0.4 to 0.8)	0.3 (0.02 to 0.6)
Wheat	0.9 (0.7 to 1.1)	0.4 (0.3 to 0.6)	0.5 (0.2 to 0.8)
Soy	0.5 (0.4 to 0.7)	0.1 (0.1 to 0.3)	0.4 (0.2 to 0.6)
Other	3.5 (3.1 to 4.0)	3.3 (2.9 to 3.8)	0.2 (−0.4 to 0.8)
Any	9.3 (8.7 to 10.1)	7.1 (6.4 to 7.8)	2.3 (1.3 to 3.2)

TABLE II. Weighted probable* prevalence estimates of individual FAs

	All		
Allergen	S2S (n = 15,322)	SPAACE (n = 15,022†)	Difference
Peanut	1.2 (0.9 to 1.5)	0.9 (0.7 to 1.2)	0.3 (−0.1 to 0.6)
Tree nut	1.4 (1.1 to 1.7)	1.2 (0.9 to 1.5)	0.2 (−0.2 to 0.6)
Fish	0.6 (0.4 to 0.8)	0.6 (0.4 to 0.9)	0.0 (−0.3 to 0.3)
Shellfish	1.1 (0.9 to 1.4)	1.5 (1.2 to 1.8)	−0.4 (−0.8 to −0.02)
Sesame	0.2 (0.1 to 0.3)	0.2 (0.1 to 0.3)	0.0 (−0.1 to 0.1)
Milk	1.1 (0.9 to 1.4)	0.3 (0.2 to 0.5)	0.8 (0.5 to 1.1)
Egg	0.8 (0.6 to 1.0)	0.6 (0.4 to 0.8)	0.2 (−0.1 to 0.5)
Wheat	0.4 (0.3 to 0.6)	0.2 (0.1 to 0.4)	0.2 (0.02 to 0.4)
Soy	0.3 (0.2 to 0.5)	0.1 (0.1 to 0.2)	0.2 (0.01 to 0.3)
Any	6.1 (5.4 to 6.8)	5.9 (5.2 to 6.7)	0.2 (−0.8 to 1.2)

How likely this could be a temporal *trend*?

PREVALENCE, *Odds* & INCIDENCE

Assuming a closed, stationary population (no immigration or emigration then:

$$Prevalence \cong Incidence \times Duration$$

$$\left(\frac{Prevalence\ Proportion}{1 - Prevalence\ Proportion}\right) = IncidenceRate \times Average\ Duration(\overline{D})$$

$$Prevalence\ Odds = IR \times \overline{D}$$

$$Prevalence = \left(\frac{IR \times \overline{D}}{1 + IR \times \overline{D}}\right)$$

Lash, Timothy, L. et al. Modern Epidemiology. (4th Edition). 2020. (pg. 71-74)

Prevalence, Incidence and Mean Duration

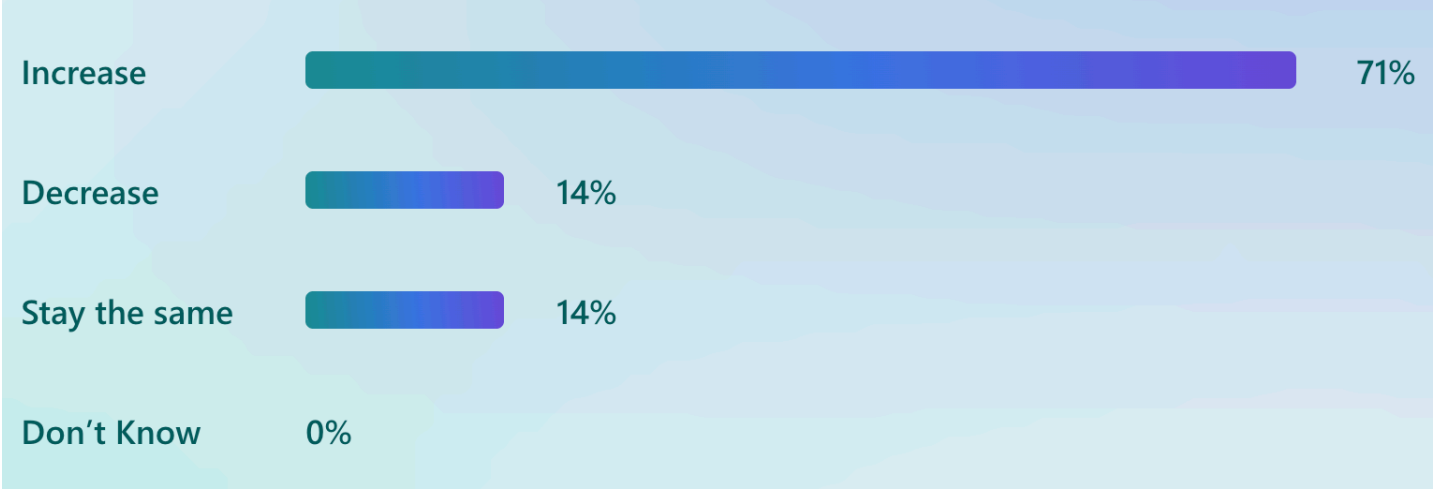
If prevalence is small (say < 0.1) then:

$$Prev \cong Incidence\ Rate \times Average\ Duration$$

"Generally, these formulas have limited practical utility because for many disease we stratify by age and there will always be immigration/emigration for age-specific prevalences when people age into and out of the age band".

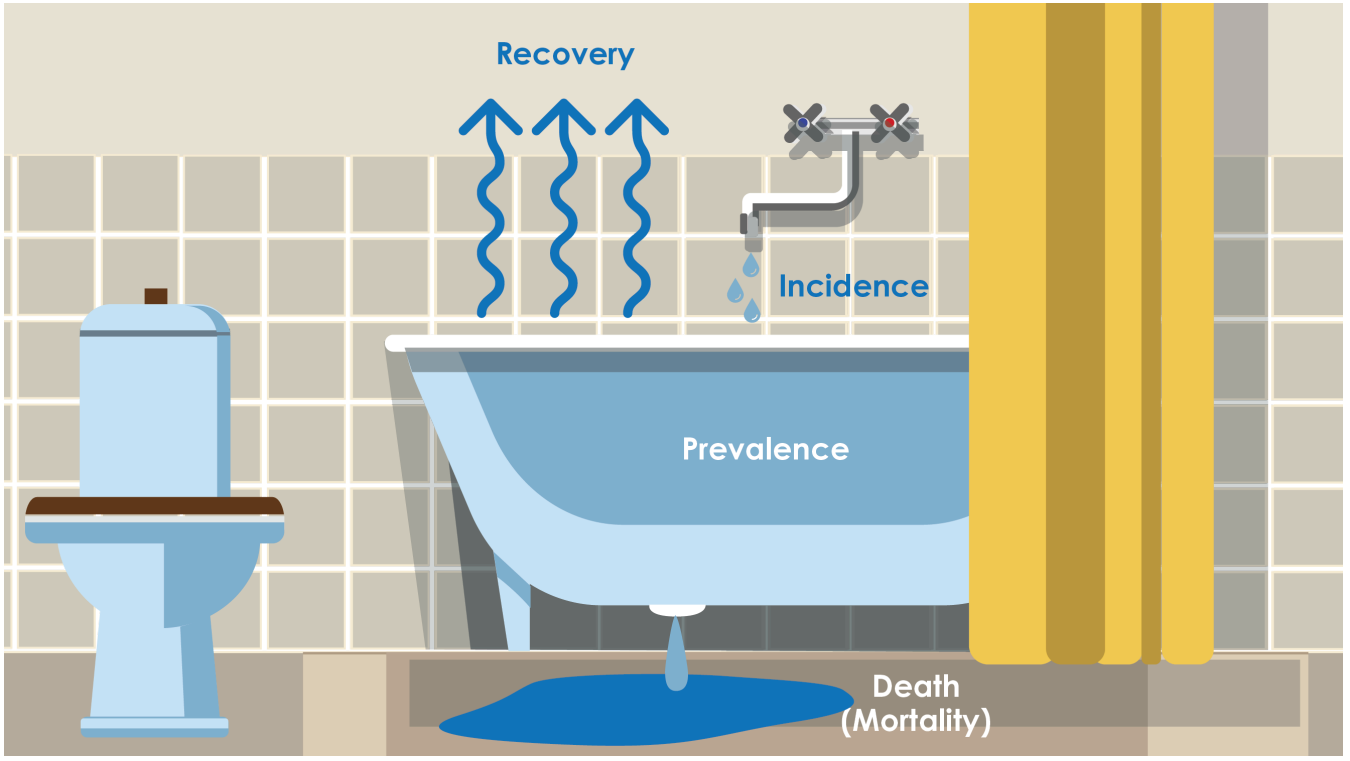
Quiz

"*Triple antiviral therapy has dramatically improved survival among patients with human immunodeficiency virus (HIV) disease. If the incidence of HIV were to remain constant, what is the expected impact of widespread triple antiviral therapy on the prevalence of HIV in the population?*"



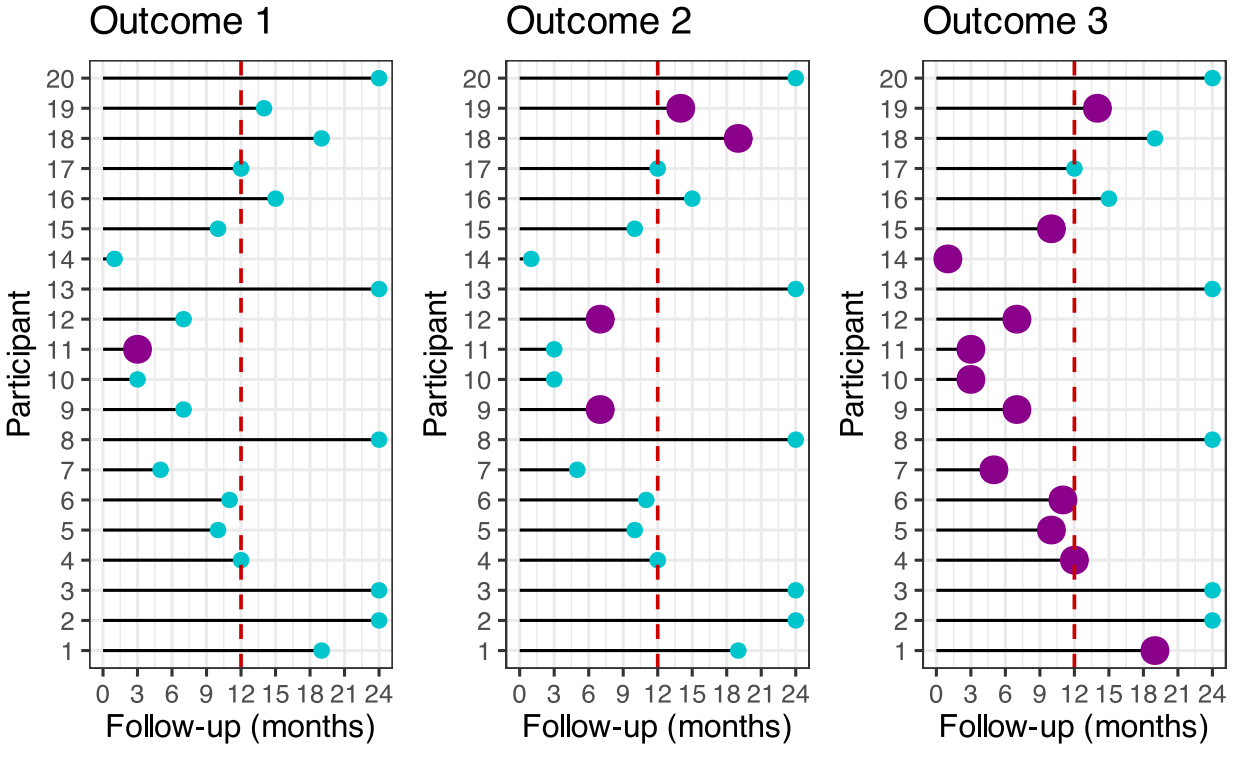
Incidence vs Prevalence

New/existing Outcomes the same?



https://www.technologynetworks.com/immunology/articles/incidence-vs-prevalence-329073

Quiz



- Cumulative Incidence /IP at 12 months? Cumulative Incidence /IP at 24 months?
- Incidence Rates? Incidence Density?
- Prevalence? Odds?

Quiz

Data	Outcomes	IP	IR	ID	Prev (12m) [1]	Odds [2]	Odds [3]
O1	1	0.05	0.04 p-y	0.04p-y	0.09	0.11	0.05
O2	4	0.2	0.16 p-y	0.18p-y	0.18	0.82	0.25
O3	12	0.6	0.48 p-y	0.54p-y	0.91	3	1.5

[1] Using the number of events (individuals with the outcome) among the population **within or during** first 12 months of follow-up.

[2] Using as reference the probabilities used to simulate the data (i.e., 0.10, 0.45, 0.75)

[3] Using the IP as the "observed" probabilities (i.e., 0.05, 0.2, 0.6)

QUESTIONS?

COMMENTS?

RECOMMENDATIONS?