

**University of Guelph
Numeracy Project**

Incidence and Prevalence



TABLE OF CONTENTS

Incidence and Prevalence.....	1
What is INCIDENCE?	1
What is PREVALENCE?	1
Numerator/Denominator	1
Introduction.....	1
The Numerator and Denominator	1
Challenges in Calculating the Numerator	2
Challenges in Calculating the Denominator	3
Incidence/Incidence Rates	3
Definition of Incidence Rate.....	3
Person-Time Incidence Rate	4
Cumulative Incidence Rate.....	5
Summary.....	6
Prevalence/Prevalence Rates	6
Definition of Prevalence Rate.....	6
Point Prevalence Rate	7
Period Prevalence Rate	7
Lifetime Prevalence	8
Summary	8
Relationship of Incidence and Prevalence	8
Introduction: Bath Model.....	8
Choice of Incidence or Prevalence	9
Summary.....	9
Glossary	10
References.....	11

Incidence and Prevalence

What is INCIDENCE?

- Incidence refers to the calculated risk of acquiring a new state (e.g. becoming sick) within a certain timeframe, which is best represented as a rate with a denominator.

What is PREVALENCE?

- Prevalence refers to a sum of instances of a particular disease within the given population at a certain point in time. It can be written as a proportion of the form:
total instances of disease in the population / total number of individuals in the population.

Numerator/Denominator

Introduction

- In epidemiology, the main goal is to investigate the relationship between probability of disease and factors which influence that probability. This investigation is typically conducted through mathematical means.

The Numerator and Denominator

- To determine the probability of disease, epidemiology requires timing, location, and other data of interest to be recorded as precisely as possible.
- *The Numerator:* Represents the number of individuals with the risk factor.
- *The Denominator:* Represents the size of the population from which the individuals were drawn.
- $\text{numerator} / \text{denominator} = \text{rate} / \text{proportion}$

- In epidemiology, when referring to the term "rate," one typically means a ratio with the numerator and denominator representing independent quantities (e.g. births/population time). Similarly, when referring to "proportion," one typically means a ratio where the numerator is related to the denominator (e.g. obesity due to one cause/obesity due to all causes).
- In considering incidence, the incidence of a disease deals with instances not related to a denominator, meaning it is not a rate.
- The incidence of disease, along with related risk factors, is continually in flux. This fact can be attributed to constant changes within the environments in which the populations reside. As a consequence, it is more helpful to consider incidence *rates* than pure incidence, as it provides a context for the incidence.

Challenges in Calculating the Numerator

- In epidemiology, it is critical to collect data as precisely as possible. The number of individuals having the risk factor is difficult to determine, as epidemiologists are not generally diagnosticians. This results in issues pertaining to confidence in the diagnoses. Of course, 100% accuracy can never be ensured, but confidence can be increased by presenting evidence of the reliability of the diagnosis (e.g. knowing the person making the diagnosis holds an MD).
- An additional issue in calculating the numerator of an incidence rate is in determining whether the case is new or old and whether the individual is a member of the study's "at risk" group.
- Finally, hospital case registers pose a hurdle in obtaining accurate data for the numerator. Some cases are clear (e.g. stab wounds); however, others may represent rare diagnoses (e.g. insomnia), which can be misdiagnosed (e.g. anemia). Due to this effect, known as the "Iceberg Effect," reported numbers of rare diseases may not remotely represent the actual total.

Challenges in Calculating the Denominator

- As with the numerator, it is critical that the denominator represent accurate data. This information should represent the population from which "at risk" individuals have been drawn.
- A crucial element in ascertaining the correctness of the denominator is to fully understand the health issue being studied. As you probably realize, the definition of the disease generally drives the choice of denominator.
- Let us consider the example of investigating the incidence rate of Alzheimer's Disease. The denominator must represent the population from which "at risk" individuals are being drawn. Here, it makes sense to ignore children, adolescents and young adults, as these groups are not known to suffer from this disease. With that said, it does not make sense to exclude any sex or ethnic group, unless your study is specifically interested in particular subsets of those groups.

Incidence/Incidence Rates

Definition of Incidence Rate

- An incidence rate refers to the regularity of new instances of an occurrence within an "at risk" population during a certain time period.

Note, this is not the same as incidence, which is not related to a denominator.

- In epidemiology, incidence rates are usually applied to:
 - Disease register studies
 - Cohort studies
 - Trials
- Considering our previous discussion of the numerator and denominator, disease incidence rates use both, but these rates also require a time period and location.

- Usually, there are two types of incidence rates employed:

- person-time incidence rate
- cumulative incidence rate

Person-Time Incidence Rate

- also known as “incidence rate,” with a range from zero to infinity.
- When referring to the numerator/ denominator concept, as it relates to the person-time denominator, it represents the length of time the study population has been "at risk." The resulting equation is given below:

$$\text{Person-time incidence rate} = \frac{\text{New instances over a particular time period}}{\text{Length of time the population has been "at risk" during this time period}}$$

- Below are some disadvantages to the above equation:
 - An assumption is made that the disease will be evenly distributed over the time period.
 - A constant rate is assumed, but it is known that many diseases pose differing risks at different ages and times.
- Accounting for the aforementioned disadvantages, the person-time incidence rate has been re-expressed in the following equation:

$$\frac{\text{New instances within a certain time period}}{\text{"At risk" population within that time period}}$$

The fraction produced is typically converted to an appropriate form (e.g. percentage, rate per thousand, etc.).

Cumulative Incidence Rate

- also known as “incidence rate,” which often results in confusion.
- often used interchangeably with risk, with a rate from 0 to 1 (or 0 to 100 percent).
- Referring back to the previous equation:

$$\frac{\text{New instances within a certain time period}}{\text{"At risk" population within that time period}}$$

The cumulative incidence rate is used when the baseline population is placed in the denominator. This rate represents a calculation of the probability of a new instance happening within a population, which explains why it is often interchanged with the term "risk."

- For this equation, the person-based denominator is based on a fixed time, generally in the first-half of the study. For instance, in a year-long study, the denominator could be the baseline population.
- In researching a recurring disease, using the cumulative incidence rate, the researcher cares about all (not only new) instances. It is correct, in this circumstance, to use a baseline "person" denominator. The rationale is that the entire population will still be "at risk."

Summary

- The chart provided below makes the assumption that the numerator (i.e. number of new cases) is identical for each measure:

Person denominator (cumulative incidence rate)	Person-time denominator (person-time incidence rate)
<ul style="list-style-type: none">• Ranges from 0 to 1• Measures absolute risk (probability) of new disease (eg cases/10000 people)• Can be used to construct relative risk• Incidence rates can be calculated with population estimates (eg from a census or disease register)• Can only be used directly in cohort studies where study participants are enrolled around the same time	<ul style="list-style-type: none">• Ranges from 0 to infinity• Not clearly interpreted as a measure of absolute risk (eg 50 cases per 1000 person years)• Can be used to construct relative risk• Person-time cannot be calculated in population estimates• Can be used either when enrolment is at about the same time OR when enrolment is spread over time.

Prevalence/Prevalence Rates

Definition of Prevalence Rate

- Some feel that prevalence is more a proportion than it is a rate. Dissimilar to incidence, those individuals having already had the disease do not need to be removed from the denominator. So, mathematically, prevalence is a proportion (ranging from 0 to 1 or 0 to 100 percent).
- In epidemiology, prevalence rates are generally found in the following:
 - Cross-sectional studies
 - Disease register studies

- We will discuss the following 3 types of prevalence rates:

- 1) Point prevalence rate
- 2) Period prevalence rate
- 3) Lifetime prevalence rate

Point Prevalence Rate

- The point prevalence rate represents all instances of a disease in a particular location at a particular point in time.
- Point Prevalence Rate =
$$\frac{\text{Every instance of the investigated disease at a particular time}}{\text{"At risk" population at the same particular time}}$$
- ** Recall, the resulting value will range from 0 to 1 (or 0 to 100 percent).
- A disadvantage to this method is that it does not always consider the disease in a way that is helpful to organizations that consider longer-term effects.

Period Prevalence Rate

- Period prevalence attempts to overcome the pitfalls of prevalence studies taken at a single point in time. All instances, whether new, old or recurring, are counted over a specified period.
- Period Prevalence Rate =
$$\frac{\text{Every instance of the investigated disease within a defined time period}}{\text{Average "at risk" population during the time period}}$$
- Period prevalence is a combination of the ideas of incidence and point prevalence, and can be helpful in investigating recurring diseases (e.g. depression). Each of incidence and point prevalence on their own have a tendency to negatively misjudge the import of such issues, making a combination of the two a better estimate.

Lifetime Prevalence

- Lifetime prevalence is a superior expansion on the concept of period prevalence. It is a representation of the proportion of the population that have ever had the disease.
- Lifetime Prevalence =
$$\frac{\text{Number of individuals having ever had the disease during their lifetime}}{\text{"At risk" population (from beginning of time period)}}$$
- The majority of information for lifetime prevalence rates is generally collected from birth cohort studies, where individuals are followed from birth. Measures of lifetime prevalence usually permit one to see the commonality of particular diseases in the frequency of their occurrence (e.g. the proportion of the population ever having had asthma could be measured at the age of 5, 10, 15, 30, 45, etc. years, until death).

Summary

- Incidence and prevalence rates are similar in the ease of use of their respective equations; however, the collection of data can be challenging. In order to calculate prevalence rates, one needs a true case definition, a means of collecting data related to the numerator, and to define and calculate a suitable denominator. Each of these requirements are comparable to those of incidence rates.

Relationship of Incidence and Prevalence

Introduction: Bath Model

- If we consider incidence and prevalence as being related to a bathtub, the incidence would be represented by the water flowing into the tub, with the prevalence being depicted by the water in the tub (pool of cases). The prevalence pool is affected by individuals leaving the population (e.g. death), which can be viewed as the drainage from the tub. The water supply itself would be individuals coming into the population (e.g. births).

Choice of Incidence or Prevalence

- Generally, when studying reasons for disease, one is most likely to use the incidence rate. This is due to the fact that the incidence rate is not affected by changes in the method of care or by case-fatality, making differences between populations simpler to explain. For research on the weight of disease over short periods (e.g. headache), the incidence rate is also most likely to be used, as point prevalence would underestimate the issue. This is due to the fact that the point prevalence value would overlook the survivors and deaths.
- When considering the weight of chronic diseases, even those that are rare (e.g. ALS), one is most likely to use the prevalence rate. For studies of disease risk factors, health activities, and even causes of disease, the prevalence rate is the favoured gauge. Many epidemiologists often see the prevalence rate as a substandard measure in comparison to the incidence rate; however, this is not the case, as each measure has both pros and cons, making their respective value dependent on the issue under investigation.

Summary

- Rate of disease is a gauge of the amount of disease (i.e. numerator) in relation to the population size (i.e. denominator).
- Incidence and prevalence are the dominant rates in epidemiology, distinguishable by their respective foci. In other words, incidence is concerned solely with new events, while prevalence considers both new and old cases. These measures are used to produce a number of views on health and disease.

Glossary

- Incidence:** Incidence refers to the calculated risk of acquiring a new state (e.g. becoming sick) within a certain timeframe, which is best represented as a rate with a denominator.
- Incidence rate:** refers to the regularity of new instances of an occurrence within an "at risk" population during a certain time period.
- Prevalence:** refers to a sum of instances of a particular disease within the given population at a certain point in time. It can be written as a proportion of the form: $\frac{\text{total instances of disease in the population}}{\text{total number of individuals in the population}}$.
- Prevalence rate:** refers to the regularity of all instances of an occurrence within an "at risk" population, which is commonly written as a proportion.

References

Bhopal, R. 2002. *Concepts of Epidemiology*. New York: Oxford University Press Inc.