Slide 35

Population-time =

1. the sum of the size of the population at risk during each small time segment

or

2. the average size of the population at risk multiplied by the length of the time interval

In practice, we estimate population time in one of two ways.

If we know the size of the population at risk during each small segment of time, then we can add up these sizes to obtain the total amount of population time.

Alternatively, if we can estimate the average size of the population we can (often more easily)

2. Estimate population time as the product of the average size of the population at risk and the number of time units in the interval during which cases occur.

Slide 36

Estimating population-time - method 1

Total population-time over 9 months =

200 + 199 + 198 + 197 + 195 + 194 + 193 + 192 + 190

= 1,758 person-months

= 146.5 person-years

However, cases are not at risk for a full month.

The first method is illustrated here, for our small population example. As a first approximation to the total population-time over 9 months, we can simply add up the number of people at risk (200+199+198+197+195+194+193+192+190), which is 1,758 person-months, or 146.5 person-years. This procedure assumes that once a person becomes a case s/he is no longer at risk to become another case), since we have removed that person from the population at risk for subsequent months.

That is a reasonable assumption for diseases like coronary heart disease, for example. The disease is regarded as having occurred one time, even if different manifestations occur at various times thereafter. One shortcoming of the above calculation, though, is that it ignores the fact that a person who becomes a case is no longer at risk for the rest of that month, either.

Slide 37

Estimating population-time - method 1 - better

Total population-time over 9 months =

199.5 + 198.5 + 197.5 + 196 + 194.5 + 193.5 + 192.5 + 191 + 189

= 1,752 person-months

= 146 person-years

assuming that cases develop, on average, in the middle of the month

5/20/2002 Incidence and prevalence 37

We can improve our approximation by adjusting for the fact that once a person becomes a case, s/he is no longer at risk for the rest of the month. If we suppose that cases develop randomly during the months, it is approximately equivalent to treat them as though they occurred in the middle of the month. In that case, each person who becomes a case is at risk for a half-month, on average.

So the population time at risk during the first month is 200 people less 0.5 for the one person who became a case, equals 199.5. The result is the same as counting the 199 people who did not become cases and adding 0.5 for the person who did become a case. Making this adjustment to the person-time for each month gives us an estimate that is smaller by the number of cases times 0.5 months.

Slide 38

Estimating population-time - method 2

Average size of the population at risk during the 9 months = 195.3 (1,758 / 9) or approximately: (200 + 188) /2 = 194

Population-time = 195.3 x 9 months or (approximately) 194 x 9 months

= 1,746 person-months

= 145.5 person-years

5/20/2002 Incidence and prevalence 38

If we do not know the number of people at risk in each small interval during the follow-up period, or if it is too tedious or not worth the effort to take the sum, we can approximate the average size of the population at risk by taking the average of the size at the beginning and the size at the end, and then multiplying this average size by the length of the time period. This method assumes that the number of cases ia about the same in each month.

Slide 39

Equivalent to - method 2

Take initial size of population at risk and reduce it for time the people were not at risk due to acquiring the disease:

200 - 12/2 = 194 (approximately)

Population-time = 194 x 9 months

= 1,746 person-months

= 145.5 person-years

An equivalent way to estimate the average population size is to take the starting population and subtract half of the number of cases, since if we assume that cases occurred evenly during the period, each case was at risk for, on average, half of the period. So we take the 200 people we started with and subtract a half-person for each of the 12 cases. The result (194) is the approximate population size, which we then multiply by the length of the time period.

All of these methods of estimating population-time assume that people are no longer at risk once they become a case. For some events we may want to regard people as still at risk of experiencing another event (e.g., experiencing a minor fall). In that case people could be regarded as being at risk during the entire period, so we would not need to reduce the size of the population at risk for the number of cases. But for various reasons we might want to define the event as "a first fall", in which case we would proceed as before.

Slide 40

Incidence rate ("incidence density")

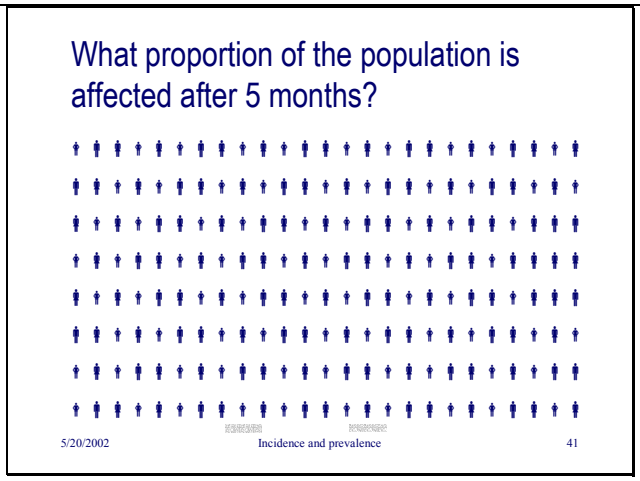
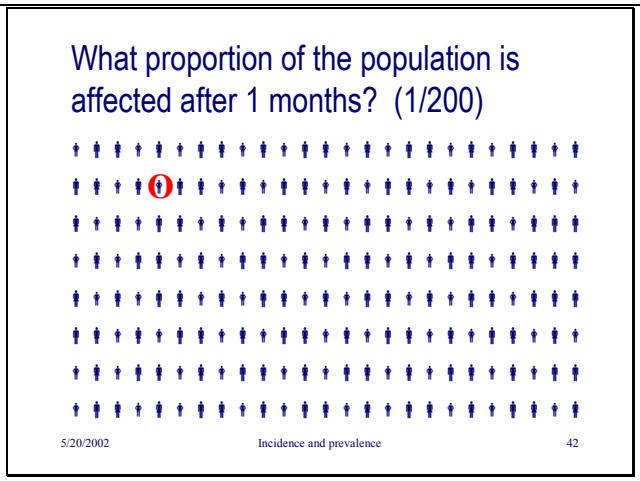
$$\frac{\text{Number of new cases}}{\text{Avg population at risk} \times \text{Time interval}}$$

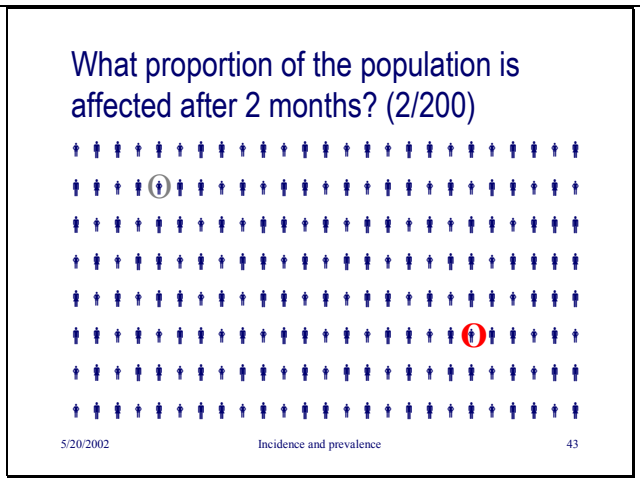
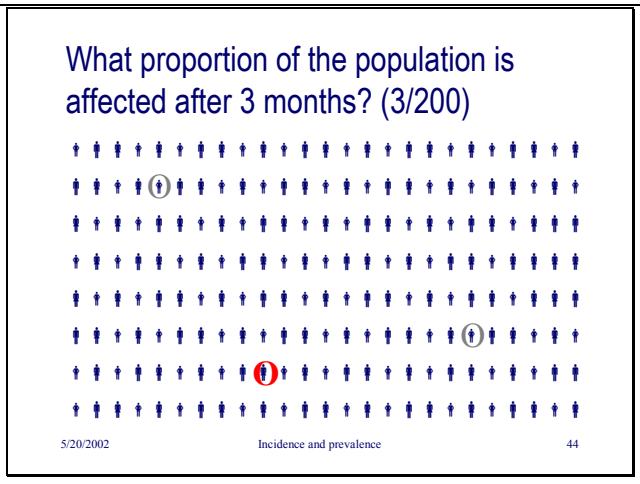
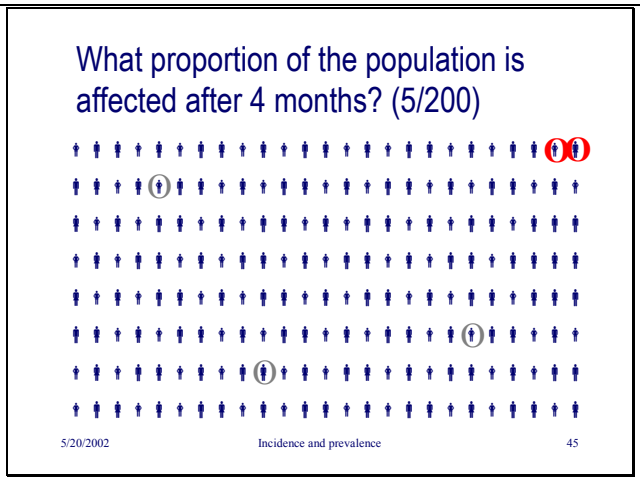
$$= \frac{\text{Number of new cases}}{\text{Population-time}}$$

So our (average) incidence rate, sometimes called "incidence density" is defined as the number of new cases divided by population-time (the average size of the population at risk multiplied by the length of the time interval).

Why have two terms (incidence rate, incidence density) for the same measure? You might suppose that the reason is so we can avoid repeating a word if we need to mention the measure twice in the same paragraph (I have heard that English is one of the richest languages in the world in that way).

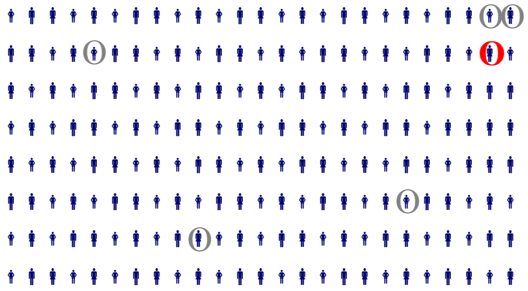
The problem is that the terminology used in epidemiology varies by person, place, and time. Different authors use different terms, different schools use

		different terms, and the terms in use change over time. Ideally the field will standardize on a set of terms, but until that happens we need to know the synonyms. So “incidence rate” (the newer term) and “incidence density” (which according to one author is not actually a “density”) both refer to the same measure.
Slide 41		What if we ask a slightly different question (i.e., we change our objective slightly)? Instead of asking for the rate of occurrence, let us ask what proportion of the population is affected after a certain time interval, say 5 months.
Slide 42		In month 1, there was one case.

Slide 43		In month 2, a second case occurred.
Slide 44		After month 3, the case count was 3.
Slide 45		Two cases occurred in month 4, bringing the count to 5.

Slide 46

6 / 200 = 0.03 = 3% = 30 / 1,000
in 5 months



5/20/2002 Incidence and prevalence 46

And in month 5 one more case occurred, so that at the end of 5 months the proportion of the population that was affected was 6 cases in 200 people, or 0.03. If we prefer not to use decimals, or at least would rather have fewer decimal digits, we can express this proportion as a percentage (3%), which is the same as "per 100", or we can express it per 1,000, as 30 per 1,000. This the same proportion, and it is a matter of personal preference (or sometimes convention) how we express it.

It is important, though, that we indicate the location of the decimal point, so that if we wish to write 0.03 as 30, we had better write it as 30 per 1,000 or 30/1,000. It is also important that we indicate the time interval. After all, this proportion was different after each month passed, so if we don't specify "5 months" then the 0.03 is quite ambiguous.

Slide 47

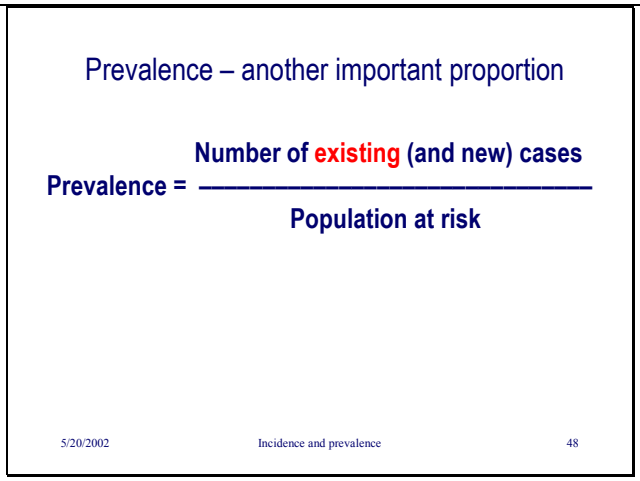
Incidence proportion ("cumulative incidence")

$$\text{5-month CI} = \frac{\text{Number of new cases}}{\text{Population at risk}}$$

5/20/2002 Incidence and prevalence 47

The measure we have just derived is called incidence proportion or cumulative incidence. It's formula is simply the number of new cases (regardless of what happened to them after they developed the disease) divided by the size of the population at risk. Here, we do not have to divide by the time interval, since we are not defining a rate of occurrence **per unit time**, but rather a summation of what has happened **over a time interval**. As noted, we **do** need to specify the length of the time interval.

Slide 48



Cumulative incidence is a **proportion**, the proportion of a population at risk that has become affected by the end of a specified time interval. Another proportion often used in epidemiology is called prevalence (and, mercifully, only that). Prevalence is the proportion of a population at risk that is affected at a given time, which could be at a given moment (which is called "point prevalence") or over a period of time ("period prevalence"). The distinction between the two is often unclear, however, and usually one just says "prevalence".

A critical distinction between cumulative incidence and prevalence is that incidence counts only newly occurring cases, whereas prevalence counts all cases that exist at the moment (or at any time during the period).