Short Report

Neuro
epidemiology

Published online: March 18, 2010

Neuroepidemiology 2010;34:267–269
DOI: 10.1159/000297758

Design and Methods of Population Surveys

Amanda G. Thrift

Baker IDI Heart and Diabetes Institute, and Department of Epidemiology and Preventive Medicine, Monash University, Melbourne, Vic., Australia

Introduction

Population surveys are undertaken on individuals to describe basic features within a defined population. The 2 main types of population surveys are prevalence studies and incidence studies. In these studies, the main emphasis is on providing reliable estimates of disease, or risk factor, prevalence and disease incidence. They may also be used to formulate hypotheses for new investigations. The accuracy of a population survey relies upon obtaining an accurate measure of both the variable being measured (numerator) and the population or sample (also known as the reference population or the denominator). This report provides a brief overview of these types of surveys.

Reference Population

In a population survey, it is essential to have an accurate figure for the population denominator. This may be required to keep account of response rates or to enable accurate incidence or prevalence rates to be calculated.

Some population surveys, such as a census, have complete enumeration of the population. This means that everybody within that population is surveyed. Because of the expense of conducting a census, governments will usually conduct these only every 5 or 10 years. In some regions, the coverage of the census is greater than 98%, and so they provide a near complete picture of the population [1].

When there are constraints on time and costs, it may be more efficient to choose a representative sample. There are many different ways to choose the sample, and the method used will depend upon the hypothesis being investigated. Choosing the sample may be done randomly or purposively [2].

Simple random sampling, such as random selection from an electoral roll, gives everyone within the population an equal chance of being chosen [2]. Other random methods of sampling include stratified, multistage and cluster sampling [2]. In stratified sampling, the population is divided into categories based on a particular characteristic, such as age. Random samples are then generated within each of these strata. Multistage sampling involves selecting samples in 2 or more stages. In the first stage, a health centre might be chosen from a list of known health centres, while in the second stage patients within each health centre are selected. This is a useful way of conducting surveys when a list of second stage units is unknown. Cluster sampling occurs when this second stage does not involve sampling. In this example, it would mean that every patient within that health centre is assessed or that every person within that health centre who has the particular characteristic is assessed. Because there are different numbers of sampling units within each cluster, there is a loss of precision in such estimates obtained [2]. The choice of sampling method used will depend upon the research question and the resources available.

Purposive sampling could be used when a certain number of respondents are required within each age and gender group. If this is not done in a random way within each age and sex category, the sample generated may not be representative in a particular characteristic of the population that has not been taken into account [2].

Counting the Characteristic of Interest (the Numerator)

It is important to obtain an accurate numerator. In a census, where everyone is being assessed, the numerator is simply the proportion of the people with the characteristic in question. In this instance, obtaining an accurate numerator is dependent upon having questions framed in such a way that reliable answers are generated. For questions about age and gender, these estimates are usually considered very reliable. Questions relating to other factors are often chosen from standardised question banks.

When counting the incidence or prevalence of a disease it is critical to have an accurate diagnosis. In neurological diseases, the patient's signs and symptoms do not always lead directly to the diagnosis of the disease [3]. Instead the signs and symptoms indicate the site of the lesion within the nervous system. Assigning a specific cause of the disease may then be done based on the results of diagnostic tests and with the aid of a skilled diagnostician [3]. As an example, patients presenting with rapidly developing clinical signs of focal disturbance of cerebral function may have suffered either an ischaemic stroke or an intracerebral haemorrhage [4]. Identifying which of these diseases is present will require the use of neuroimaging with or without the use of other diagnostic tests [5].

Prevalence Studies

Prevalence studies are conducted to identify every single person who has a particular disease or characteristic within a community, including old and new cases. They are usually conducted in a cross-sectional manner, using methods such as population sampling and door-to-door surveys. Sometimes prevalence is estimated from cohort studies or by indirect calculation from inci-

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

0251–5350/10/0344–0267$26.00/0

Accessible online at:
www.karger.com/ned

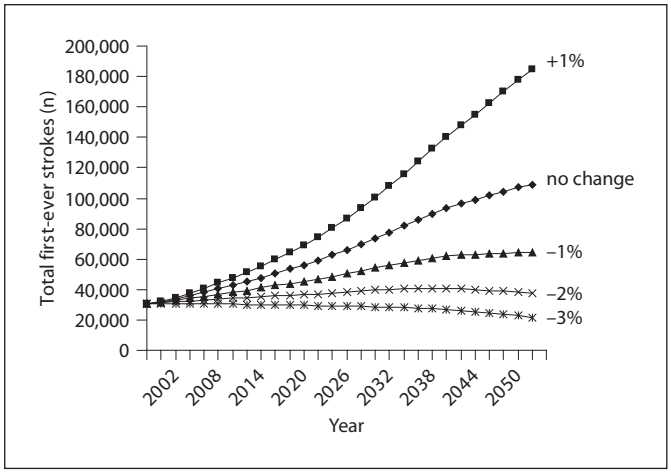

Fig. 1. Changes in the number of first-ever strokes in Australia (1997–2050). Estimates obtained using incidence rates obtained from the North East Melbourne Stroke Incidence Study and predicted projections of the Australian population [11, 14]. Percentages relate to the annual change in incidence rate.

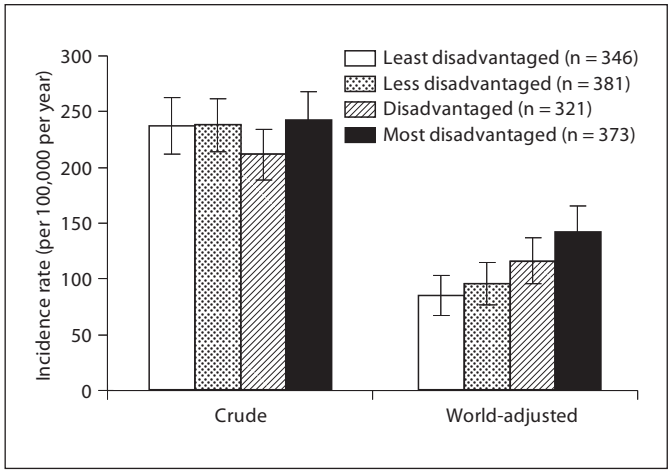

Fig. 2. Incidence of stroke according to socioeconomic group in the North East Melbourne Stroke Incidence Study (rates are both unadjusted and adjusted to the world population of Segi) [11, 12].

dence studies where cases have been followed up at regular intervals over time.

Accurate estimates of the prevalence of a disease assist health administrators to plan health services. They also assist in formulating strategies for preventing disease occurrence, particularly when information on risk factors is also collected. Prevalence figures can be used to identify current and potential future gaps in health care facilities and services. In some diseases, such as stroke, a large proportion of survivors are likely to be dependent on either a family or professional caregiver for assistance in everyday activities. Identifying services that could be useful to caregivers may help them to better deliver the care their relatives require.

There are some disadvantages of prevalence studies. First, when prevalence studies are focused on a disease, only survivors of the disease can be counted. Although estimates are relatively representative of people with the condition when there is a long lag time between diagnosis of the disease and mortality from the disease, prevalence estimates under-represent people with a short course of the disease. These studies are also time-consuming to conduct, relatively expensive, and subject to recall bias and confounding [6].

Suggested criteria for the conduct of prevalence studies include the use of a standard definition of the disease (or risk factor), and a population that is large, well-defined and stable [7].

Incidence Studies

Incidence studies enable one to measure the size of the disease problem. The aim of these studies is to find every single new case within the identified population over a defined period of time. Depending on the questions being asked, these studies can be limited to particular age and sex groups. For example, in stroke, some investigators have limited the study population to those over the age of 25 years [8]. This is because very few events occur in those younger than this age.

Apart from providing information on disease burden, incidence rates can be used to make projections of future burden of the disease. For example, when the age and sex of the projected Australian population was taken into account, it was projected that the overall number of strokes occurring would approximately triple in a 50-year period if incidence rates remained unchanged (fig. 1).

Incidence studies have a number of advantages. First, they overcome the limitation of mortality studies by including non-fatal occurrences of the disease and not just those dying from the condition. These studies can also be used to assess geographical, seasonal and secular variations; economic costs to patients, caregivers and society; and when patients are followed up over time they can be used to assess both short- and long-term prognosis. Furthermore, when the same methods are used to repeat the study in the same population, time trends in incidence can be assessed. Data on the cause-specific trends in incidence provide important feedback on public health measures being utilised in the community.

In some diseases, criteria for conducting incidence studies have been proposed [9, 10]. Many of these criteria relate to the complete ascertainment of cases. When studies in different regions have been conducted using the same criteria, then comparing incidence rates may provide insights into the development of the disease. When different methods are used for conducting these studies, it will remain unclear whether actual differences exist or whether it is the differing methods used that actually explain any differences in disease incidence.

Standardisation of Incidence and Prevalence Rates

Diseases usually occur in different rates in different age groups. When this is the case, the disease is said to be confounded by age. When a disease is more frequent in older ages, then a community with larger numbers of older people will have a greater crude stroke incidence of the disease than a community com-

prised mainly of younger individuals. This greater incidence will be apparent even if the risk of developing the disease is the same in both regions. Age-adjustment is a statistical process that removes the differences in the age composition between populations and enables them to be compared independent of their age structure. As an example, figure 2 shows the incidence of stroke according to different socioeconomic groups in the North East Melbourne Stroke Incidence Study [11]. The crude incidence is similar between regions. However, the least disadvantaged groups are older and, when incidence is adjusted to the world population of Segi, it is evident that the incidence of stroke is actually greater in more disadvantaged regions [12].

Age-adjustment should be calculated using the appropriate standard population. Comparisons between European studies should be undertaken using the standard European population, while worldwide comparisons should be undertaken using a world population comparator [12, 13].

Summary

Population surveys are used to determine the burden of disease. Accurate estimates of disease incidence and prevalence require accurate assessment of the population denominator and the disease numerator. Determining the population denominator may be relatively simple when census data for the region are available, but when samples are used this may require more care by the investigators. Random samples tend to be more representative than purposive samples. To obtain a true estimate of the disease numerator, care must be taken to ascertain all cases of the disease and its diagnosis should be carefully defined. For disease incidence, complete ascertainment may require the use of multiple overlapping sources, whereas for disease prevalence a door-to-door approach may be the best option. Finally, when comparing disease prevalence or incidence over time or between countries, one must be careful to remove the confounding effects of age by age-adjusting to a standard population.

References

1 Trewin D: How Australia Takes a Census (catalogue number 2903.0). Canberra, Australian Bureau of Statistics, 2006.
2 Armitage P, Berry G, Matthews JNS: Statistical Methods in Medical Research, ed 4. Oxford, Wiley-Blackwell, 2002.
3 Kurtzke JF: Neuroepidemiology. Ann Neurol 1984;16:265–277.
4 Thrift AG, Donnan GA, McNeil JJ: The epidemiology of intracerebral hemorrhage (review). Epidemiol Rev 1995;17:361–381.
5 Caplan LR: Intracerebral haemorrhage. Lancet 1992;339:656–658.
6 Beaglehole R, Bonita R, Kjellström T: Basic Epidemiology. Geneva, World Health Organization, 1993.
7 Feigin VL, Lawes CM, Bennett DA, Anderson CS: Stroke epidemiology: a review of population-based studies of incidence, prevalence, and case-fatality in the late 20th century. Lancet Neurol 2003;2:43–53.
8 Dalal PM, Malik S, Bhattacharjee M, Trivedi ND, Vairale J, Bhat P, Deshmukh S, Khandelwal K, Mathur VD: Population-based stroke survey in Mumbai, India: incidence and 28-day case fatality. Neuroepidemiology 2008;31:254–261.
9 Sudlow CL, Warlow CP: Comparing stroke incidence worldwide: what makes studies comparable? Stroke 1996;27:550–558.
10 Feigin V, Hoorn SV: How to study stroke incidence. Lancet 2004;363:1920.
11 Thrift AG, Dewey HM, Sturm JW, Paul S, Gilligan AK, Srikanth VK, Macdonell RAL, McNeil JJ, Macleod MR, Donnan GA: Greater incidence of both fatal and non-fatal strokes in disadvantaged areas: The North East Melbourne Stroke Incidence Study. Stroke 2006;37:877–882.
12 Waterhouse J, Muir C, Correa P, Powell J: Cancer incidence in five continents. Lyon, IARC, 1976.
13 Ahmad OB, Boschi-Pinto C, Lopez AD, Murray CJL, Lozano R, Inoue M: Age standardization of rates: a new WHO world standard (GPE discussion paper series: No. 31). World Health Organization, 2000. www.who.int/healthinfo/paper31.pdf.
14 Australian Bureau of Statistics: Population Projections (2002–2101), Australia (catalogue No 3222.0). Canberra, Australian Bureau of Statistics, 2003.

Assoc. Prof. A. Thrift
Stroke Epidemiology, Baker IDI Heart and Diabetes Institute
75 Commercial Road
Melbourne, Vic. 3004 (Australia)
Tel. +61 3 8532 1100, Fax +61 3 8532 1111
E-Mail amanda.thrift@bakeridi.edu.au