

Cite this article:
Khamis AH. Is Comparison
of the Prevalence of Disease
Appropriate as Health
Indicator between
two Populations?
Stoma Edu J. 2014;1(2):
92-96.

IS COMPARISON OF THE PREVALENCE OF DISEASE APPROPRIATE AS HEALTH INDICATOR BETWEEN TWO POPULATIONS?

**Amar Hassan Khamis^{1a*},
Maanas Shah^{2b}**

1. Biostatistics and Research Unit,
Hamdan Bin Mohammed College
of Dental Medicine (HBMCDM),
Dubai, United Arab Emirates
2. Department of Periodontology,
Hamdan Bin Mohammed College
of Dental Medicine (HBMCDM),
Dubai, United Arab Emirates

a. PhD, DEA, MSc, BSc, Associate Professor
b. BDS, MSD, CAGS, Clinician

Abstract

Introduction: Prevalence is the most frequently used health indicator in order to assess the magnitude of a disease in a public health environment. It is a fraction that relates the number of screened disease cases divided by the total number of screened cases. When comparing two different study populations it is not possible to predict the proportions of old and new cases. This, when combined with variations in the health care delivery systems, makes the comparison of prevalence within two populations questionable. The objective of this study was to discuss the validity of the comparison of prevalence between two different populations.

Methodology: Mathematical derivatives were presented to express the prevalence of a disease in a given population. Further fragmentation of the equations led to various probabilities pertaining to the number of old versus new cases that contributes to the prevalence of any existing disease. These calculations were applied to a theoretical example and final confirmation of its applicability was completed using various published scenarios from the Scientific Database.

Results: The decomposition of the formula of prevalence to probabilities that measure new, old and normal case probabilities out of the screened individuals will lead to the fact that, not all the parts of these formula are comparable, due either to different settings, health systems or even to the time of exposure to a given impairment.

Conclusion: The conclusion is that the prevalence of a disease between two different populations is unlikely to be comparable.

Keyword: prevalence, prevalence comparison, health indicators, Dubai healthcare city.

Introduction

The recent definition of the word *disease* in the Oxford English dictionary is: a condition of the body or some part or organ of the body, in which its functions are disturbed or deranged; morbid physical conditions; a departure from state of health, especially when caused by structural change (1). This definition identifies a wide spectrum of the concept of the word *disease* from *influenza* to *arthritis* and from *tuberculosis* to *cancer*. In epidemiology and in biostatistics, despite the above-mentioned definition, a disease is considered as a dichotomous variable (yes or no) regardless of the status of the illness or the different states of the patient.

The *prevalence* or *prevalence proportion* is an indicator that is frequently used in epidemiology to describe the number of people with a disease in a given population. The proportion of a population found to have a condition (typically a disease) is derived by

Received: 24 September 2014
Accepted: 06 October 2014

* Corresponding author:

Associate Professor Amar Hassan Khamis,
PhD, DEA, MSc, BSc
Head of Biostatistics and Research Unit
Hamdan Bin Mohammed College of Dental
Medicine (HBMCDM)
34, Al Zahrawi Building
Dubai, PO - 505097
United Arab Emirates
Tel/Fax: +97144248787
e-mail: amar.hassan@dcdm.ac.ae

comparing the number of people found to have the condition within the total number of people studied, and is usually expressed as a fraction, as a percentage or as the number of cases per 10,000 or 100,000 people.

Point prevalence is the proportion of a population that has the condition at a specific point in time. *Period prevalence* is the proportion of a population that has the condition at some time during a given period, and includes people who already have the condition at the beginning of the study period as well as those who acquire it during that period. *Lifetime prevalence* (LTP) is the proportion of a population that at some point in their life (up to the time of assessment) have experienced the condition (2), or simply, the prevalence of a disease is the proportion of people with a given disease at a given time.

Prevalence is conventionally expressed as the proportion or percentage of cases in a given population at a specified time. Prevalence differs from, but is often confused with incidence which refers to the number of new cases arising during a specified time (3). Cases may move out of the pool from which prevalence is taken because of mediating factors such as spontaneous recovery, intervention or death. Therefore, although an increase in the incidence will usually increase the prevalence, this will not be the case if the disease is self-limiting. The prevalence also depends on the method of case finding, the health system and available sources of diagnoses and the training of the personnel. In addition, the length of time a population has been exposed to the risk of disease would also have an impact on the prevalence.

Finally, in the light of these considerations, prevalence rates are not an appropriate comparison. Changes in the incidence within a population affect the proportion of cases with a disease but are reflected in changes in incidence rates whereas prevalence remains the proportion of diseased patients.

The aim of this paper is to discuss the validity of the comparison of prevalence between two different populations.

Mathematical Derivations

We can express the prevalence of a disease as follows:

$$P = \frac{N + O}{T}$$

where

P is prevalence

N is the number of new cases

O is the number of old cases

T is the total number of screened cases

Equation (1) can also be presented in the form

$$P = \frac{N + O}{N + O + C}$$

where C is the number of cases free from the disease and $T = N + O + C$

Thus, the probability of new cases arising is N/T and the probability of old cases being recorded is O/T . Combined, these two probabilities make up the prevalence (See Appendix). As long as detection methods are similar, new cases are comparable between two different populations whereas the existing (old) cases are of unknown duration and may represent disparate incidences or disease exposure and are therefore not comparable. Combining old and new cases into an overall prevalence makes inappropriate the direct comparison of two populations.

There follows a theoretical example to illustrate this idea and three scenarios taken from the medical literature in order to highlight the issues involved.

Theoretical Example

Let us suppose we have two populations, A and B. Their prevalence for the disease was evaluated twice with a delay of 12 years (2000 and 2012). The following table shows the outcome:

Year	Size	2000	2012
Population A	1000	1.5%	2.5%
Population B	1000	2.0%	2.5%

Table 1. Prevalence of the disease in the theoretical example

The prevalence in the two populations in 2012 is equal, but if we see the history it seems that the evolution of the disease prevalence in the two populations is different. When we see the history, the comparison becomes uncertain, because it may be not only the speed at which the disease appears in population A which is higher than that in population B, i.e. the incidence is higher than in population B, but the health system may not be as efficient in diagnosing the disease or may not be able to prolong the survival time.

SCENARIO 1 (4)

A study was formulated to compare the prevalence of systemic health conditions (SHC) between African American and Caucasian edentulous patients coming for complete dentures (CD) at an urban dental school. The authors concluded that among selected completely edentulous patients at an urban dental school, two out of three patients had at least one SHC. This exploratory study suggests there may be health status differences between African American and Caucasian patients in this setting (health system and exposure).

SCENARIO 2 (5)

The World Health Organization has evaluated the prevalence of dental caries for the 12-year-olds. Data from recent studies show that the mean caries prevalence among 12-year-olds in Latvia, Lithuania, and Estonia was 5.8, 4.9, and 4.6, respectively. Absence of caries was recorded in 5% of Latvian and Estonian and 12% of Lithuanian 12-year-olds. Fifteen-year-old Latvians and Lithuanians averaged 8.1 and 7.0 DMFT, which, owing to the absence of radiographic examination, may be a substantial underestimation of real caries levels (exposure time). The possible adverse effects of the privatization of dental care and the benefits of increased access to fluoride dentifrice in these countries have not been evaluated yet. The extremely poor oral hygiene seen in epidemiologic surveys indicates that fluoride dentifrices may not be widely used (health system). The caries levels in the Baltic states resemble those commonly encountered a couple of decades ago in the Nordic countries (exposure time).

SCENARIO 3 (9)

According to a National Oral Health Survey report 2004, caries prevalence in India was 51.9%, 53.8% and 63.1% at ages of 5, 12 and 15 years respectively in different parts of the country. Available literature from 1940-1960 shows the prevalence of dental caries in India had a varied picture. This study was carried out to measure the caries prevalence and treatment needs in school children of 6-14 years of age residing in coastal areas of West Bengal (9). The overall caries prevalence in the permanent dentition was 28.06%. It was of 25.39% in boys and of 30.86% in girls (9).

The Discussion sections mentions the comparison of these figures with other studies. A similar result was reported by Saravanan et al. in 2008. (10) They reported that the prevalence of caries in the permanent dentition was 26.5%. The prevalence of caries increases significantly with age in the permanent dentition. A very low level of dental caries was reported by (6, 7, 8).

In 12 year-old school children in urban Kenya in 1984, a caries prevalence of nearly 22% and a 0.5 DMFT were reported by the author. Low caries prevalence was reported by Knutson in 1947. He reported the caries prevalence in 6-year-old children in Nicollet County, Minnesota. The caries prevalence was 24.3%. In 1939 Sarkar examined 18,445 school children up to 16 years of age in West Bengal. He reported that 13.3% of the children had defective teeth. Furthermore, in 1931, he examined 2,000 children and found that 14.4% of them had defective teeth. Sgan-Cohen et al. found the prevalence of dental caries to be very low among 5 and 12-years old children (9).

Discussion

The decomposition of the formula of prevalence to probabilities that measure new, old and normal case

probabilities out of the screened individuals will lead to the fact that, not all the parts of these formula are comparable, due either to different settings, health systems or even the time of exposure to a given impairment.

The epidemiological studies publishing data on prevalence of a particular disease are generally observational and not experimental. These studies involve the risk of different compounding factors, known and unknown, that can affect the outcome of the investigation at the time of exposure.

These compounding factors determine the extent to which observed associations are causal. It may give rise to illicit associations when in fact there is no causal relationship, or at the other extreme, it may lead to incomprehensible effects due to a true cause.

Scenario 1 is an example of how different health systems between two populations are likely to lead to two samples which are not comparable in the sense of prevention, case detection, the delay of diagnosis and intervention. Likely, these factors play a major role over the time of exposure to the disease and will affect the probability of *old cases*, among the screened individuals.

Scenario 2 is another example of how the history of the disease in a population will affect the comparison between the prevalence in two different populations, considering this caries problem and its relation with the index of DMFT or even with cavities; both these incidences took place back in history and early discovery of such problems needs an efficient integrated delivery health system that links both the school health system and the public systems; this is another handicap that leads to the occurrence of the probability of *old cases* among screened individuals which are not comparable.

Scenario 3 highlights the fact that even within the same population, the comparison between two different prevalence for the same disease with different delays of time will raise the same problem. This time, the health system is not be responsible, but the attitude of the patients themselves towards the disease and their ability to follow the rigorous instructions given to them. Some of them will sustain and live with the disease without any desire to cure it completely, whilst others will seek care and will eliminate their disease. So within the delay we are facing a situation where the probability of *old cases* among the individuals screened is not comparable due to the lack of awareness or even to the individuals' bad attitude towards their health.

Conclusion

Comparing prevalence between two different populations or even within one population at different times is questionable due the incomparable setting of old cases included in this comparison.

Bibliography

1. Keeling MJ, Rohani P. Modeling Infectious Diseases in Humans and Animals. 1st Ed. Princeton: Princeton University Press; 2007.
2. Rothman KJ. Epidemiology: An Introduction. Oxford: Oxford University Press. p. 53. 2012.
3. Moscicki EK. The prevalence of 'incidence' is too high. ASHA. 1984;26(8):39-40.
4. Szykowska E, Kaste LM, Schreiner J, Gordon SC, Lee DJ. Comparison of systemic health conditions between African American and Caucasian complete denture patients. J Prosthodont. 2014;23(5):341-346.
5. Bjarnason S. High caries levels: problems still to be tackled. Acta Odontol Scand. 1998; 56(3):176-178.
6. Frencken JE, Sithole WD, Mwaenga R, Htoon HM, Simon E. National oral health survey Zimbabwe 1995: dental caries situation. Int Dent J 1999; 49(1): 3-9.
7. Pitts NB, Palmer JD. The dental caries experience of 5-, 12- and 14-year-old children in Great Britain. Surveys coordinated by the British Association for the Study of Community Dentistry in 1991/92, 1992/93 and 1990-91. Community Dent Health 1994; 11(1): 42-52.
8. Riordan PJ. Secular changes in treatment in a school dental service. Community Dent Health 1995; 12(4): 221-225.
9. Das D, Misra J, Mitra M, Bhattacharya B, Bagchi A. Prevalence of dental caries and treatment needs in children in coastal areas of West Bengal. Contemp Clin Dent. 2013; 4(4): 482-487.
10. Saravanan S, Kalyani V, Vijayarani MP, Jayakodi P, Felix J, Arunmozhi P, Krishnan V, Sampath Kumar P. Caries prevalence and treatment needs of rural school children in Chidambaram Taluk, Tamil Nadu, South India. Indian J Dent Res 2008; 19(3):186-190.

Appendix

Mathematical Derivations

We can express the prevalence of the disease D as follows:

$$P = \frac{N + O}{S} \quad (1)$$

where

N is the number of new cases

O is the number of old cases

T is the number of screen cases

Equation (1) can be also presented in the form

$$P = \frac{N + O}{N + O + C} \quad (2)$$

where

C is the number of individuals free from the diseases, let $T = N + O + C$

If we are dividing equation (2) by 2, we obtain

$$P = \frac{\left(\frac{N}{T}\right) + \left(\frac{O}{T}\right)}{\left(\frac{N}{T}\right) + \left(\frac{O}{T}\right) + \left(\frac{C}{T}\right)} \quad (3)$$

where

(N/T) , is the probability of new cases among screened cases;

(O/T) , is the probability of old cases among screened cases;

(C/T) , is the probability of normal cases among screened cases.

$$P_N = \frac{N}{T}$$

is the probability of new cases among the screened population; here, the utilization of the same protocol of diagnose of a given disease D will lead to comparable probabilities irrespective of the populations under study. We know exactly what is going on as the cases will be diagnosed during screening, and all the cases have no idea that they have the disease before (*Rational*). The question about delay may rise at this point, but it will not affect the comparison.

$$P_O = \frac{O}{T}$$

is the probability of the old cases among the screened population; this probability depends on the history of the disease D in a given population, the health system and the exposure time to the disease D in a population. The exposure time is a real problem, as the right time of exposure is unknown for old cases, so this probability is not comparable between two different populations.