Registers in cardiovascular epidemiology

Reitsma, J.B.

Chapter 8

Lack of standards in direct standardisation

JB Reitsma, GJ Bonsel, LJ Gunning-Schepers, JGP Tijssen

Accepted for publication in
The American Journal of Public Health

ABSTRACT

Objectives - To review the recent use of direct standardisation and determine the types of standards applied.
Methods - We retrieved all articles using direct standardisation in the American Journal of Public Health, the European Journal of Public Health, the JAMA and the BMJ in 1996 and 1997.
Results - In 19% (14/75) of all articles using direct standardisation no information about the standard was found. Only 15% (11/75) of the articles used one of the widely available standards provided by the World Health Organization. This percentage was only slightly higher, 27% (10/37), among studies using routine data sources.
Conclusions - In direct standardisation, the type of standard can influence the results. Therefore, readers need to know which standard has been applied and, to facilitate the comparison between studies using routine data sources, the use of a widely available standard is strongly advocated. At present, these guidelines are insufficiently pursued.

INTRODUCTION

Many researchers compare rates across populations or analyse trends over time. If groups or years differ with respect to factors associated with the event under study (e.g. age, gender or race), the comparison of the overall or crude rate may be misleading. One popular way to address these differences in confounding factors is direct standardisation, in which a weighted average of stratum specific rates is calculated.[1-4] Although standardised rates are not a substitute for analysing specific rates, they serve as a convenient summary measure. The purpose of determining such a standardised summary measure is in the comparison with other rates adjusted to the same standard. Readers therefore need to know which standard has been applied. The use of a widely available standard (European or World population standard)[5] is strongly advocated in studies involving routine data sources (like vital and hospital statistics data, and data from cancer registers). The use of widely available standards enables other researchers to adjust their rates to the same standard, which facilitates comparison and interpretation of differences. We reviewed the recent use of direct standardisation in the medical literature, in particular the choice and specification of the standard applied.

METHODS

The formula for calculating standardised rates is given by:

$$SR = \frac{\sum (r_i \times w_i)}{\sum w_i}$$

where SR is the standardised rate, r_i is the specific rate in stratum i, and w_i is the weighting factor for stratum i.

We manually searched all 1996 and 1997 issues of two general medical journals (BMJ and JAMA) and two public health journals (American and European Journal of Public Health) for articles using direct standardisation. For each article we determined the type of standard used in the standardisation procedure and classified it into one of the following five categories: (1) unknown standard, if no information was given; (2) sample standard, if the unknown composition of one of the analysed groups was used; (3) specific standard, the standard was specified, but no reference to retrieve the actual composition; (4) specific standard with reference given or weights cited; (5) one of the widely available standards provided by the *World Health Organization* (WHO).[5]

Results

We retrieved 75 articles using direct standardisation while searching all 1996 and 1997 issues of the four journals. In 19% (14/75) of the articles using direct standardisation we were unable to identify the standard that was applied (table 1). Only 15% (11/75) of the articles used one of the widely available standards of the WHO. Surprisingly, this percentage was only slightly higher, 27% (10/37), among studies using routine data sources.

Discussion

Direct standardisation addresses differences in confounding factors by calculating a weighted average of specific rates.[1-4] The absolute height of a standardised rate is of minor importance and varies with different weights (w_i). It is the *hypothetical* number of events that would have occurred if the population structure had been identical to that of the standard. The rationale for standardised rates is in the comparison with other rates adjusted to the same standard.

Table 1. Standards applied in 75 articles using direct standardisation in two general medical and two public health journals during 1996 and 1997. The number of studies involving routine data sources[#] rather than study-specific data is given between brackets.

Characteristic	Am J Public Health	Eur J Public Health	JAMA	BMJ	Total
No. of studies using direct standardisation	29 (14)	8 (6)	10 (2)	28 (15)	75 (37)
Type of standard used:					
- widely available standard[†]	1 (1)	3 (3)	0	7 (6)	11 (10)
- specific standard with weights given or cited	5 (2)	1 (0)	3 (0)	2 (1)	11 (3)
- specific standard, but without reference[‡]	10 (8)	1 (1)	4 (2)	3 (3)	18 (14)
- sample standard[§]	8 (0)	1 (0)	2 (0)	10 (0)	21 (0)
- unknown standard	5 (3)	2 (2)	1 (0)	6 (5)	14 (10)

[#] routine data sources include mortality data, hospital statistics data or data from cancer registers.

[†] European or World population standards provided by the WHO.[5]

[‡] for instance, adjusted to the age structure of Wales in 1990, but no reference to retrieve the actual composition of the standard.

[§] the overall distribution within the particular cohort itself has been used as standard, without further specification.

Using direct standardisation does not eliminate the effect of a confounding factor, but merely keeps it constant. Standardisation works best when the effect is constant over all strata. If all specific rates (r_i) between two populations show the same ratio, then the two standardised rates will show that same ratio, regardless of the weights applied. If all specific rates are higher in one population then the standardised rates will be higher in that population, irrespective of the choice of standard. However, if some specific rates are higher in one population and some higher in the other then the choice of standard will determine which population will have the higher rate after adjustment. Therefore, if specific rates in the groups being compared do not bear consistent relations across strata, in other words interaction is present, any summarisation procedure will yield interpretation problems. In these situations, the choice of standard can affect the results.[6,7] Doll and Cook showed that the ranking of countries based on the age standardised incidence of cancer varied depending on the choice of standard.[6] On the other hand, Spiegelman and Marks performed similar studies in which they showed that the choice of standard had little effect in the comparison of the directly standardised mortality figures.[7]

Each standard has its own advantages and disadvantages that may guide your choice.[1,4,8] The use of one of the widely available standards as provided by the WHO (European or World population standards) is strongly recommended to enable researchers from other countries to adjust their rates to the same and easy accessible standard. If applicable, the use of a similar weighting scheme for men and women is endorsed as readers often compare rates between men and women. Unfortunately, during some time the WHO provided European and World standards with different weights for men and women.[9]

In conclusion, standardised rates are no substitute for analysing specific rates. As the choice of standard may affect the comparison, readers need to know which standard has been applied. In one fifth of all articles using direct standardisation we could not retrieve that information. Widely available standards, to facilitate the comparison between studies using routine data sources, were infrequently used.

ACKNOWLEDGEMENT

This work was supported by a grant from the Netherlands Heart Foundation (grant no. 42.012).

References

1. Kitagawa EM. Standardized comparisons in population research. Demography 1964;1:296-315.
2. Anderson S, Auquier A, Hauck WW, Oakes D, Vandaele W, Weisberg HI. Standardization and stratification. In: Statistical methods for comparative studies. Techniques for bias reduction. New York: John Wiley & Sons, 1980:113-139.
3. Breslow NE, Day NE. Rates and rate standardization. In: Statistical methods in cancer research. Volume II: the design and analysis of cohort studies. New York: Oxford University Press, 1987:48-79.
4. Hoem JM. Statistical analysis of a multiplicative model and its application to the standardization of vital rates: a review. Int Stat Rev 1987;55:119-152.
5. World Health Organization. World Health Statistics Annual 1995. Genova: WHO, 1996: XXV.
6. Doll R, Cook P. Summarizing indices for comparison of cancer incidence data. Int J Cancer 1967;2:269-279.
7. Spiegelman M, Marks HH. Empirical testing of standards for the age adjustment of death rates by the direct method. Human Biology 1966;38:280-292.
8. Kitagawa EM. Theoretical considerations in the selection of a mortality index, and some empirical comparisons. Human Biology 1966;3:309-317.
9. World Health Organization. World Health Statistics Annual 1992. Genova: WHO, 1993: XXII.