VALUE IN HEALTH 17 (2014) 629–633

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval

An Updated Method for Risk Adjustment in Outcomes Research

Suhail A.R. Doi, FRCP, PhD, Jan J. Barendregt, PhD, Chalapati Rao, PhD*

School of Population Health, University of Queensland, Brisbane, Australia

ABSTRACT

Objective: To demonstrate why meta-analytic methods need modification before they can be used to aggregate rates or effect sizes in outcomes research, under the constraint of no common underlying effect or rate. **Methods:** Studies are presented that require different types of risk adjustment. First, we demonstrate using rates that external risk adjustment through standardization can be achieved using modified meta-analytic methods, but only with a model that allows input of user-defined weights. Next, we extend these observations to internal risk adjustment of comparative effect sizes. **Results:** We show that this procedure produces identical results to conventional age standardization if a rate is being standardized for age. We also demonstrate that risk adjustment of effect sizes can be achieved with this modified method but cannot be done using standard meta-analysis. **Conclusions:** We conclude that this method allows risk adjustment to be performed in situations in which currently the fixed- or random-effects methods of meta-analysis are inappropriately used. The latter should be avoided when the underlying aim is risk adjustment rather than meta-analysis.

Keywords: age standardization, burden of disease, fixed-effects model, meta-analysis, population standardization, quality-effects model, random-effects model.

Introduction

The term *standardization* refers to the process of facilitating comparison of summary measures of burden or risk of disease across populations. Such standardization can be done in two ways. *Internal* standardization refers to the process of ensuring that summary measures adequately reflect the distribution of burden or risk of disease within subpopulations of the same overall population. If the subpopulations are investigated independently of each other (e.g., in separate studies), the overall summary measure for the total population needs to take account of the actual size of each subpopulation within the overall population structure so that the overall summary reflects the actual population meaningfully. For example, a summary measure of mortality reported in different subpopulations by age can be combined into a summary measure for the total population after standardizing against the actual distribution of ages. When such standardized rates are compared across different populations, they can be interpreted as the mortality rate for an average member of each specific population. Thus, mortality rates in Australia versus India, obtained from different subpopulations, need to be internally standardized against the actual population structure in Australia and India, respectively, if a summary for each country is to be compared.

External standardization can also be done by replacing the internal standard described above with a common external standard against which subpopulations from different populations are standardized, thus removing the subpopulation effect completely. This is of particular importance to studies of quality improvement, in which many estimates of disease burden or risk are strongly dependent on subpopulation, with rates of incidence or mortality being much higher or lower between different subpopulations. In this situation, the differences between populations independent of the confounding by sizes of subpopulations with different risks can be determined by standardizing against a common external standard. In this sort of standardization, the standardized rate is in itself useful only for comparison and has no intrinsic interpretation.

The process of *risk adjustment* encompasses both standardization and other procedures for accounting for the effects of subpopulations with different risks. In this article, risk adjustment and standardization will both refer to methods of adjustment based on weighted averages in which the weights are chosen to provide an "appropriate" basis for the comparison (i.e., a "standard"). The latter is generally either the subpopulation sizes from each of the populations in the comparison or from a relevant external population. A common method in epidemiology for this purpose has been direct standardization because it can be applied on the basis of any subpopulation distribution, for example, on the basis of age, geographical clusters, and cancer incidence. Direct standardization is simply a process of weighted averaging of the subpopulation-specific rates to arrive at a standardized estimate that reflects a given subpopulation structure. The distribution of the "standard" provides the weights and usually represents the *current* or *most common* subpopulation structure for internal and external standardization, respectively,

* *Address correspondence to:* Suhail A.R. Doi, Clinical Epidemiology Unit, School of Population Health, University of Queensland, Herston Road, QLD 4006, Brisbane, Australia.

E-mail: sardoi@gmx.net; s.doi@sph.uq.edu.au.

1098-3015$36.00 – see front matter

http://dx.doi.org/10.1016/j.jval.2014.05.003

and could represent subpopulation sizes based on age or geographic cluster or any other distribution of whatever standard is to be applied. This provides, for each population, one risk-adjusted or standardized rate that reflects the appropriate contribution of the subpopulation-specific risk or rates to the standard. In this article, we demonstrate, using two examples, which risk adjustment through direct standardization can be achieved using modified meta-analytic methods, but only with our model [1] that allows input of user-defined weights. The advantage here is that this method can now be extended to any standard and any effect size (ES) other than rates.

Methods

A modification was undertaken of the quality-effects model [1] of meta-analysis that allows moving the model from meta-analysis to risk adjustment. This model uses a risk of bias weighting scheme in addition to inverse variance weighting and the modification entailed removing inverse variance weights and replacing bias weights with normalized subpopulation weights from a standard population. The subpopulation weights are applied using a modification of our bias adjustment procedure in meta-analysis [1,2] for each subpopulation to come up with a weighted average that represents the single risk-adjusted or standardized estimate across the subpopulations. This weighted averaging procedure *does not* use inverse variance weights and thus is *not* a meta-analysis. Therefore, if subpopulation rates are being combined, it would give an equivalent result to direct standardization used in epidemiology. We do not use log-transformed rates because back transformation would result in pooled estimates that depart from those computed using the standard method. The standard method used for the computation of the directly standardized rate (DSR) is given by:

$$\mathrm{DSR}=\frac{1}{\sum_{i=1}^{k} W_j}\times\sum_{j=1}^{k}\frac{w_j o_i}{n_j} \tag{1}$$

where O_j is the observed number of events in subpopulation (age group) j, n_j is the number of individuals in subpopulation (age group) j (or the population × person-years at risk), and w_j is the weight based on the number/total (proportion) of individuals in the age-group subpopulation j. The computations for the variance and confidence intervals of this DSR are outlined in Table 1.

However, this estimate can be derived by using a different procedure. If weights are given by $w_j^{\alpha}=Q_j+\hat{\tau}_j$ (see Table 1 for the computation of $\hat{\tau}_j$), $Q_j = N_j/N_{\max}$, N_j is the subpopulation size, and ES_j is the subpopulation effect estimate of interest, which could be an ES, rate, or proportion, the directly standardized effect estimate (DSE) is given by

$$\mathrm{DSE}=\frac{\sum\left(w_j^{\alpha}\times \mathrm{ES}_j\right)}{\sum w_j^{\alpha}} \tag{2}$$

In the computation of this DSE, rates can also be one of the effect estimates standardized and in this special case, zero rates are imputed to have variances based on a single observed event as a continuity correction (see Table 1). The same method can be used by substituting ES_j for any other ES. For the odds ratios, however, careful consideration should be given to whether the marginal or the conditional odds ratios are of interest in a particular analysis, given the mathematical fact that the marginal and conditional odds ratios are nonequivalent [3]. Two examples are given below of the application of this procedure to risk adjustment in outcomes research. In the first example, external risk adjustment is done via the new procedure and compared with the direct method of age standardization to demonstrate equivalence. In the second example, this is then extended to internal risk adjustment of a relative risk (RR) measure using the incidence rates of cancer in each subpopulation as the weights and demonstrates how this may be extended beyond risk adjustment for rates.

A simulation was also run (for example 1) under sampling variability by allowing $[O_j \sim \mathrm{Poisson}(O_j)$ after replacing any $O_j=0$ with 1. Thus, within each of the 18 age-group subpopulations, O_j was now generated from a Poisson distribution with mean O_j. A thousand iterations of each set of rates were run using Ersatz version 1.3 (Epigear International Pty Ltd., Brisbane, Australia). Coverage of the confidence interval and percent bias was then computed as described by Burton et al. [4].

Results—Some Examples of Risk Adjustment

Example 1: External Risk Adjustment Across Age Groups and a Simulation Study

Individual death records with multiple cause of death were the primary source of data and accessed through the Australian Bureau of Statistics for the period 1999 and 2006. Deaths were coded according to the *International Statistical Classification of Diseases, 10th Revision* by using the automated Mortality Medical Data System and results have been reported previously [5]. The Australian population age distribution in 2006 was used as the external standard population for the purpose of risk adjustment. To examine mortality trends and differentials across time, we had created three estimates of a risk-adjusted mortality rate from renal failure due to diabetes using standard methods as follows [5]: 1) Risk-adjusted rates (underlying cause rate) for diabetic renal disease based on deaths coded to diabetic nephropathy; 2) Risk-adjusted rates (multiple cause rate 1) for diabetic renal disease based on 1) above and additional deaths coded to diabetes without complications but with renal failure as a multiple cause; and 3) Risk-adjusted rates (multiple cause rate 2) for diabetic renal disease based on 1) and 2) above and additional deaths coded to diabetes with other complications (except nephropathy) but with renal failure as a multiple cause.

The risk-adjusted cause-of-death rate of patients via our new procedure is obtained as follows: 1) compute the cause rate of each age subgroup of patients; 2) create a standardized weight ($N_j/N_{\max}$) from the age composition of the external standard population adopted as the 2006 population in our case; and 3) apply the weighting procedure above to obtain the age-standardized rate. Table 2 depicts the standard computation versus the modified meta-analytic procedure results. The pooled rates are identical because the process in both cases is weighted averaging. The confidence intervals differ marginally even though the process for risk adjustment here is completely different from the standard computation for direct standardization of rates.

Results of the simulation are also shown in Table 2 and demonstrate excellent coverage of the confidence interval by this method and reaffirm the appropriateness of the use of the normal approximation to the Poisson distribution for the generation of the variance of the subpopulation rates even when there are low event rates. Also, as expected with empirical weights, the estimator is biased because of covariance between the effect and the weights. This was demonstrated in the case of rates, however, to be a very small percentage of the magnitude of the effect and can therefore be ignored (Table 2).