LETTER TO THE EDITOR

Recommendation on unbiased estimation of population attributable fraction calculated in "prevalence and risk factors of active pulmonary tuberculosis among elderly people in China: a population based cross-sectional study"

Ahmad Khosravi[1] and Mohammad Ali Mansournia[2*]

Abstract

Population attributable fraction (PAF) refers to the proportion of all cases with a particular outcome in a population that could be prevented by eliminating a specific exposure. The authors of a recent paper evaluated the prevalence and estimated the PAFs for risk factors of TB among elderly people in China [Inf Dis Poverty. 2019;8:7]. Confounding is inevitable in observational studies and Levin's formula is of limited use in practice for unbiasedly estimating PAF. In a complex survey design, an unbiased estimation of the PAF can be calculated using a sample-weighted version of the Miettinen formula or a sample weighed parametric g-formula. With respect to causal interpretation of PAF in public health setting, computation of PAF is logical and practical when the exposure is amenable to intervention.

Keywords: Population attributable fraction, Confounding, Sample-weighted parametric g-formula

Multilingual abstracts

Please see Additional file 1 for translations of the abstract into the five official working languages of the United Nations.
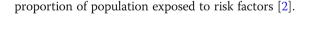
To the Editor.

We read with great interest a recent article titled [1]: "Prevalence and risk factors of active pulmonary tuberculosis among elderly people in China: a population based cross-sectional study". The authors evaluated the prevalence and identify the risk factors of TB among elderly people in China using a cross-sectional study. However, there are several concerns in the analysis.

i). In the statistical analysis section it was indicated that population attributable fraction (PAF) of each adjusted risk factor was estimated using Levin's

formula where RR is the risk ratio and p_e means proportion of population exposed to risk factors [2].

$$PAF = \frac{p_e(RR-1)}{p_e(RR-1) + 1} \qquad (1)$$

In the study, the adjusted odds ratio (OR) was used in place of RR.

PAF refers to the proportion of all cases with a particular outcome in a population that could be prevented by eliminating a specific exposure [3]. Formula 1 is unbiased in the absence of confounding and effect modification [3, 4]. Observational studies are subject to confounding which will lead to bias if Levin's formula is inappropriately applied to estimate PAFs [3]. The Levin's formula is valid only for unadjusted risk ratio [3–5]. The bias from this error will depend on the degree of confounding [6]. For a

* Correspondence: mansournia_ma@yahoo.com
[2]Department of Epidemiology and Biostatistics, Tehran University of Medical Sciences, Tehran, Iran
Full list of author information is available at the end of the article

dichotomous exposure an unbiased estimation of PAF can be calculated using the Miettinen's formula [7].

$$PAF = \frac{p_c\left(RR_{adj}-1\right)}{RR_{adj}} \qquad (2)$$

Where RR_{adj} is the adjusted risk ratio and p_c is the prevalence of exposure among the cases. This produces valid estimate in the presence of confounding, assuming exposure status and confounders are accurately measured and adjusted for.

As an example, in this study [1] the adjusted *OR* and prevalence of diabetes in the active TB cases was reported 1.83 (1.08–3.10) and 16/193, respectively. The PAF using formula 2 is 3.76% which is less than the reported value (5.52%).

ii). The term "attributable" refers to a causal interpretation [3]. One of the main assumptions underlying the PAF is no bias in study design. Therefore, the application of formula 2 in cohort design is acceptable but for case-control and cross-sectional studies, it needs more considerations. In a cross-sectional study, reverse causality and prevalence-incidence bias are the main concerns for assessing the effect of the exposure on the outcome.

iii). Another potential source of bias in the study is failure to adjust observed estimates of the prevalence of TB and exposure to risk factors for the complex sampling design employed. With such a design, the population prevalence should be adjusted using inverse probability weighting (IPW) so that the reported prevalence is appropriately adjusted for multistage and disproportionate sampling [8]. Further, the authors do not mention whether or how clustering was taken account of in the multivariable logistic regression modeling.

iv). For complex survey designs, it is necessary to adjust PAFs for the complex sampling design [9, 10]. PAF can be computed as a sample-weighted version of the Miettinen or Bruzzi formula (formula 2 in reference [9] or formula 3 in reference [10] or sample-weighted model-based standardization, also known as parametric g-formula (formula 3 in reference [9] or formula 4 in reference [10].

v). As PAF is a function of the prevalence of exposure, self-reported measurement of exposures in this study can be lead to bias in the estimation of PAF. As reported in this study, the self-reported and local health documentation search of diabetes was not sufficient to estimate the real distribution.

vi). With respect to causal interpretation of PAF in a public health setting, computation of PAF is logical and practical when the exposure is amenable to

intervention [6]. Therefore, it is less apparent why the attributable fraction for unmodifiable risk factors such as age and sex may be of use.

In sum, unbiased estimation of PAF requires several assumptions which are often ignored in practice. We recommend using sample-weighted version of Miettinen formula or sample weighed parametric g-formula [3, 11].

Additional file

Additional file 1: Multilingual abstracts in the five official working languages of the United Nations. (PDF 485 kb)

Author details
[1]Department of Epidemiology, School of Public Health, Shahroud University of Medical Sciences, Shahroud, Iran. [2]Department of Epidemiology and Biostatistics, Tehran University of Medical Sciences, Tehran, Iran.

References
1. Zhang CY, Zhao F, Xia YY, Yu YL, Shen X, Lu W, et al. Prevalence and risk factors of active pulmonary tuberculosis among elderly people in China: a population based cross-sectional study. Infect Dis Poverty. 2019;8(1):7.
2. Darrow LA. Commentary: errors in estimating adjusted attributable fractions. Epidemiology. 2014;25(6):917–8.
3. Mansournia MA, Altman DG. Population attributable fraction. BMJ. 2018; 360:k757.
4. Darrow LA, Steenland NK. Confounding and bias in the attributable fraction. Epidemiology. 2011;22(1):53–8.
5. Flegal KM. Bias in calculation of attributable fractions using relative risks from nonsmokers only. Epidemiology. 2014;25(6):913–6.
6. Rockhill B, Newman B, Weinberg C. Use and misuse of population attributable fractions. Am J Public Health. 1998;88(1):15–9.
7. Miettinen OS. Proportion of disease caused or prevented by a given exposure, trait or intervention. Am J Epidemiol. 1974;99(5):325–32.
8. Mansournia MA, Altman DG. Inverse probability weighting. BMJ. 2016; 352:i189.

9.   Heeringa SG, Berglund PA, West BT, Mellipilan ER, Portier K. Attributable fraction estimation from complex sample survey data. Ann Epidemiol. 2015; 25(3):174–8.
10. Graubard BI, Fears TR. Standard errors for attributable risk for simple and complex sample designs. Biometrics. 2005;61(3):847–55.
11. Mansournia MA, Etminan M, Danaei G, Kaufman JS, Collins G. Handling time varying confounding in observational research. BMJ. 2017;359:j4587.