Author's response to reviews

Title: Methodologies used to estimate tobacco-attributable mortality. A review.

Authors:

   Monica Perez-Rios (mrmonica@usc.es)
   Agustin Montes (mrmonte@usc.es)

Version: 2 Date: 22 June 2007

Author's response to reviews: see over
Methodologies used to estimate tobacco-attributable mortality. A review.

Authors: Pérez-Ríos, M$^{1,2}$ MPH; Montes, A$^2$ MD, PhD.

Authors’ affiliations
(1) Department of Epidemiology. Directorate-General for Public Health. Galician Regional Health Authority. Santiago de Compostela. Spain
(2) Department of Preventive Medicine and Public Health. University of Santiago de Compostela. Santiago de Compostela. Spain

Mónica Pérez Ríos

Address for correspondence and reprint requests
Tel: +34-981-540044. Fax: +34-981-572282

E-mail: mrmonica@usc.es

Abstract (words) 195
Text (words) 4605
References (number) 121
Tables (number) 2
Figures (number) 0
Article type Review
ABSTRACT

Objective

One of the most important measures for ascertaining the impact of tobacco on a population is the estimation of the mortality attributable to its use. To measure this, a number of indirect methods of quantification are available, yet there is no consensus as to which furnishes the best information. This study sought to provide a critical overview of the various different methods of attribution of mortality due to tobacco use.

Method

A search was made in the Medline database until March 2005 in order to obtain papers that addressed the methodology employed for attributing mortality to tobacco use. A search was made in the Medline database until March 2005.

Results

The different methods were obtained by means of a bibliographic review. Of the total of 7 methods obtained, the most widely used was proportional attribution of the prevalence methods, followed by the approach proposed by Peto et al. and cols., with the remainder being used in a minority of studies.

Conclusions

Different methodologies are used for estimation of tobacco attributable mortality, but their methodological foundations are quite similar in rather all of them. Mainly, they are based on the calculation of proportional attributable fractions. All methods show limitations of one type or another, in some cases specific, and in others, common to all, sometimes common to all methods and some specific.
Key words: epidemiologic methods, attributable mortality, tobacco.
INTRODUCTION

Since the association between tobacco and mortality was first discovered (1, 2), the task of attributing a given number of deaths to smoking has been and continues to be a complicated controversial process, beset by limitations and questioned from different quarters, including the powerful world tobacco industry. With the appearance of the successive revisions of the International Classification of Diseases (ICD), there has been considerable progress in the process of categorizing mortality, but very little progress methods for attributing this mortality to risk factors such as tobacco. Having roughly accurate reliable estimates of the impact of tobacco on mortality would enable facilitate to have a clearer picture to be formed of the reality of the problem caused by smoking and would be on help in the planning of health policy action measures.

The task of quantifying smoking-attributable mortality is one to which indirect methods are fundamentally applied has been performed mainly through indirect methods. This review sought to list and describes the different methods of estimating mortality attributed to tobacco use, to indicate the principal methodological differences existing among them, and to identify the possible sources of variability in the results obtained.
METHODS

In order to obtain papers that addressed the methodology employed for attributing mortality to tobacco use, a search was made in the Medline database until March 2005, using the terms, mortality, attribut,* method* and tobacco or smoke*. The search was completed with a manual review of the bibliographic references cited by the papers retrieved and of other publications, such as the monographs published by the Centers for Disease Control and Prevention (CDC). [Suggested by Rev.4] The main inclusion criteria was the use of an epidemiological method to estimate attributable mortality. Papers describing mortality, such as cohort follow-up or mortality studies were excluded unless an epidemiological analysis had been used.

Animal studies and communications presented at congresses were excluded from the search. [Suggested by Rev 4] The estimation of attributable mortality is also estimated applied to other risk factors in addition to tobacco such as alcohol consumption or obesity. In order to avoid the exclusion of valid methodologies the search was repeated without restricting it to tobacco or smoke.
RESULTS

The search yielded a total of 372 papers. Of these, 74 were finally included, as the rest did not apply mortality attribution methods. Some papers included more than one method. [Suggested by Rev 4] The unrestricted search, without the terms tobacco or smoke, did not furnish any new alternative methodology.

Perusal of these 74 papers enabled us to distinguish 2 types of mortality attribution procedures for the specific case of tobacco. The first one is based on individual analysis of deaths to ascertain what, if any, implication tobacco use may have had in them. [Suggested by Rev 2] Only three studies have applied this procedure (3-5).

The second is based on the application of indirect methods and constitutes usually the most commonly used methodology of choice in attribution of mortality. The total number of papers that employed this indirect methodology was 73, with 61 of these being yielded by the automatic and 12 by the manual search.

Seven indirect methods for estimating tobacco-attributable mortality were identified (no alternative methods were furnished by papers dealing with estimation of mortality attributable to different causes). Due to the methodology applied, the applied methodology in these 7 methods can be classified under: it is possible to join them into four categories: Relative risk-prevalence models—Prevalence-based analysis (Proportional attribution method, hybrid method by Lam—Prevalence-based analysis in cohort studies, prevalence-based analysis in case-control studies and the basic method), Peto and colleagues’ method, methodology—methodologies based on the calculation of excess mortality (Garfinkel’s and Roger’s method) and predictive models.
The methods differed from one another in terms of calculation processes, information requirements, data sources and assumptions required for their application. A summary of these methods is showed in Table 1. The main features characteristics of the different indirect methods used to estimate smoking-attributable mortality are described below.
Relative risk-prevalence models (Prevalence-based analysis).

Prevalence-based analysis or prevalence risks. Relative risk-prevalence models are based on the different distributions of the risk of dying from various tobacco-related diseases by reference in relation to the prevalence of tobacco use consumption in the population. To apply these methods, one needs it is necessary to know the prevalence of smoking in the study population, the total number of deaths due to diseases causally related to tobacco use, and a measure that summarizes the increased risk of dying due to these causes among smokers and sometimes among ex-smokers.

Attributable deaths are calculated for each cause of mortality, using the following formula:

\[ AM = OM \times PAF; \]

where AM is mortality attributed to tobacco, OM observed mortality, and PAF the population attributable fraction.

Modifications in calculation procedures because of data source let us to distinguish three methods:

- Proportional attribution method (Prevalence-based analysis in cohort studies)

This method is the most widely employed in the literature consulted (4, 6-51), and The CDC’s SAMMEC (Smoking-Attributable Mortality, Morbidity, and Economic Cost) computer software application (52) and the SimSmoke simulation model (15-17, 19) both use this methodology (53). Estimations from SAMMEC are different of the rest of methods because it includes attributed deaths to the fires and to secondhand smoke.
Attributable deaths are calculated for each cause of mortality, using the following formula:

\[ AM = OM \times PAF; \]  

where AM is mortality attributed to tobacco, OM observed mortality, and PAF the population attributable fraction.

To calculate the PAF, different methods of calculation exist, though the most widely used in this particular instance, is based on the formula proposed by Levin, which divides the population into various categories according to tobacco use (non-smokers, ex-smokers and smokers):

\[ PAF = \frac{\left(p_0 + p_1RR_1 + p_2RR_2\right) - 1}{p_0 + p_1RR_1 + p_2RR_2}; \]

where \( p_0 \), \( p_1 \) and \( p_2 \) represent the prevalence of non-smokers, smokers and ex-smokers, respectively. \( RR_1 \) and \( RR_2 \) refer to the respective risks of dying for any cause in smokers and ex-smokers, respectively compared to a baseline population of non-smokers, for any given disease.

Data are drawn from registries in the case of observed mortality, and from surveys in the case of smoking prevalence. The relative risks (RRs) employed in the calculations are extracted from a variety of sources, including: the prospective cohort study conducted by the American Cancer Society, i.e., the Cancer Prevention Study II (CPS II) with follow-up at 4 and 6 years (4, 6, 8, 17, 19, 20, 22, 24, 26, 27, 30, 31, 33, 35, 36, 38, 39, 41, 44, 45, 47, 49, 50, 59, 60); the performance of a meta-analysis (18, 21, 42, 48, 59, 61); health-mortality studies (10, 29, 37), (7, 23, 31, 32, 43); or, the average of the results yielded by the different studies (25, 34, 40, 46). A modification of this method was proposed in the 1992 Surgeon General’s report “Smoking and Health in the Americas” (57). The authors created an index for measuring the smoking maturity in a population, based on a comparison of lung
cancer rates. This index is multiplied by the disease-specific PAF to obtain an adjusted
disease-specific PAF for a country.

[Suggested by Rev 1]The CDC’s SAMMEC (Smoking-Attributable Mortality, Morbidity, and Economic Cost) computer software application (58)(52) uses applies this methodology. SAMMEC is a software is a computer application package commonly used in the United States to estimate smoking-attributable mortality due to smoking, years of potential life lost and indirect mortality costs. SAMMEC computes calculates PAF automatically after the user includes prevalence of tobacco consumption. Furthermore, the user also must supply the number of deaths, by 5-years age groups from 35 or older, for each smoking-related diagnosis. Estimations from SAMMEC can include attributed deaths to fires and to second-hand smoke. The Simsmoke model, a model used to uses this computer application to estimate deaths attributable to smoking.

Apart from being employed for calculating attributing mortality due to tobacco use, this method has also been used for estimating mortality associated with exposure to environmental tobacco smoke (59-61)(63-65), alcohol intake (8, 18, 21, 24, 29, 62-65)(8, 18, 21, 24, 29, 66-69), illicit drugs (18, 21), obesity (66, 67)(70-71), oral contraceptive use (68)(72), hypertension status (69)(73), cardiovascular processes (70)(74), and diabetes status (71)(75).

Prevalence-based analysis in case-control studies Hybrid method by Lam
Employing a similar calculation procedure to the previous method, this one emerged as a consequence of the objections raised by certain researchers when it came to using RRs to estimate smoking attributable mortality from other countries with longer histories of tobacco use (72)(76). The first country in which it was applied was China. This method has been used to estimate mortality attributable to tobacco use (73-75) in China when the epidemic was still in the initial phase.

To be able to apply this method, it is necessary to know the total of deaths for all causes among subjects aged 35 years and more for among, generally, the over-35 age group in a population during a given period of time. By interviewing survivors contemporaries, information is collected retrospectively on such decedents so as to be able to reconstruct their individual smoking patterns years before their death, smoking habits of the deceased subjects 15 years before their death. Based on a case-control study, risks are calculated estimated for smokers versus non-smokers.

Once these risks have been obtained, the population attributable fraction (PAF) can then be calculated by applying the formula:

\[ PAF = P \times (1 - \frac{1}{RR}) \]

where P is the proportion of deaths occurring among smokers and RR the relative risk calculated as OR after completion of a case-control study.

When the PAF has been calculated, deaths attributed to tobacco use (AM) in the study population can be estimated as follows:

\[ AM = OM \times PAF \]

This method has been used to estimate mortality attributable to tobacco use (77-79) and environmental exposure to tobacco smoke (80). [Suggested by Rev 4] in China.
Basic model

The Basic model (76,81) was originally applied in the context of occupational cohort studies, for the purpose of assessing confounding generated by tobacco use. This model has been employed in only one study (76,81) to estimate non-tobacco-attributable lung cancer mortality rates. Unlike the previous methods, different processes are specified here for calculating the RR of lung cancer in smokers and ex-smokers versus non-smokers. From a paper previously published (77,82) authors adapted two functions to compute rate ratios. Both of them take into account duration and intensity of smoking.

Lung cancer rate not attributable to smoking (I_o) can be calculated as follows:

\[ I_o = \frac{I}{P_0 + P_1RR_1 + P_2RR_2} \]

where I is the overall lung cancer mortality rate.

[Suggested by Rev 4] and P_0, P_1, and P_2 represent the prevalence of non-smokers, smokers and ex-smokers, respectively. RR_1 and RR_2 refer to the respective lung cancer risks of dying in smokers and ex-smokers versus non-smokers.

b) Method proposed by Peto et al.

Although this method could be defined as a prevalence-risk model, particularities in its calculation procedure and assumptions would rather be classified individually. Peto et al. (78, 79, 83, 84) established a method for estimating tobacco-related mortality in which the need for data, especially for lung cancer estimates, is less demanding pronounced than in any of the other procedures reviewed. These authors
postulate that lung cancer mortality is an indicator of the maturity of the smoking epidemic in a population, and thus, that tobacco-attributable mortality can thus be estimated by reference to the magnitude of the former by lung cancer mortality. This model enables mortality to be estimated without the need to know independently of the prevalence of smoking in the study population.

To be able to apply this method, one needs to know: the age- and sex-specific lung cancer mortality rates in the target country \( C_{\text{LC}} \) and also in never-smokers of the same population \( N_{\text{LC}} \), the relative risks for all the various diseases and disorders causally related to tobacco use, except lung cancer; and the cause-specific lung cancer mortality rates in smokers \( S'_{\text{LC}} \) and never-smokers \( N'_{\text{LC}} \), taken from a cohort study. Peto and co-workers used data drawn from the CPS II.

The calculation of the estimated tobacco-attributable mortality has two well-defined procedures: the first being applied to estimating attributed lung cancer mortality, and the second to estimating mortality attributable to all the remaining pathologies having diseases with an established causal relationship \( (55, 56)(57, 58) \).

The sex- and age-specific proportions of lung cancer deaths attributable to tobacco are obtained after applying the following formula:

\[
\frac{(C_{\text{LC}} - N'_{\text{LC}})}{C_{\text{LC}}}
\]

For the remainder of the diseases causally associated with tobacco use, the calculation process is different. The first step is to estimate the summarized smoking prevalence or smoking impact ratio (SIR), which summarizes the history of tobacco use in the population by age and sex. SIR was defined as population lung-cancer mortality in
excess of never-smokers relative to excess lung-cancer mortality for a known reference group of smokers, adjusted to account for differences in never-smoker lung-cancer mortality rates across populations \((80)(85)\). Smokers in the study population are converted into equivalent of smokers in the reference population \((Peto et al. 1992)\). The formula used for its calculation is as follows:

\[
SIR = \frac{C_{LC} - N_{LC}}{S^*_C - N^*_LC}
\]

This formula is used in all populations where lung cancer mortality rates among non-smokers are unknown. Where these data are available it is needed to normalize the formula \((80)(85)\). The second step of this process consists of computing the population etiologic fraction (PEF) on the basis of the previously calculated summarized prevalence (SIR) and the relative risks of dying due to the respective causes (RR), by age group and sex, as per the CPS II.

\[
PEF = SIR(RR-1)/(1+(SIR(RR-1)).
\]

To ensure that the resulting PEF was not exaggerated by excessively high RRs, Peto and colleagues adjusted the formula proposed by Levin \((54)(56)\) by replacing the 1 in the denominator by a 2.

Once the RRs from the CPS II had been re-analyzed and their robustness confirmed, the earlier reduction was viewed as excessive, and a reduction of 30% applied instead \((81)(86)\). In countries like China, where one’s own risks are available, the reduction applied is lower.
The last step in this procedure would involve applying the following formula:

\[ AM = OM \times PEF \]

in order to obtain the estimation of attributed mortality, AM, in accordance with the previously calculated PEF and the observed mortality, OM.

This method has only been applied to estimation of tobacco-attributable mortality (11, 78, 79, 81-83).

c) Excess mortality methods

Garfinkel’s method

Garfinkel’s method has been applied to estimate cancer mortality attributable to tobacco use (89-91). Cancer deaths due to smoking are calculated as the difference between observed and expected deaths in a population. To apply this method, age- and sex-specific cancer mortality rates are needed, and age- and sex-specific cancer mortality rates for non-smokers are computed on the basis of the CPS study (84). Attributable fractions due to smoking are calculated as the difference between observed and expected deaths. The expected deaths are related to the number of deaths that would occur if the whole population was formed by non-smokers. [Rev 1] To calculate the expected number of deaths, the follow-up over 12 years of the never smokers enrolled at the CPS I study was employed and death rates for cancer were computed. These rates were applied to the estimated number of person-years of exposure for non-smokers to obtain the expected number of deaths for each site. The attributable fractions calculated in this way have been found to be similar to those yielded by the CPS (85)(90). Garfinkel’s method was applied to estimate cancer mortality attributable to tobacco use (85-87).
Rogers's method

The method proposed by Rogers et al. (88)(93) is based on the calculation of excess mortality on the basis of known disease-specific mortality rates in 7 designated groups. Combines prevalence and mortality risk rates in order to offer more precise estimates of smoking attributable mortality. This calculation procedure attempts to avoid some problems related to previous methods as the use of risks derived from non-random sample of population, risks rarely adjust for confounding factors or the classification of the smoking status in crude categories without attending to the number of cigarettes smoked by former and current smokers. At first, age-specific smoking prevalence and mortality risks were estimated. The authors define 7 population groups.

The following 7 population groups could be distinguished by reference to the amount of cigarettes smoked (\(p\)) and classified by sex and age-group: non-smokers, light smokers, moderate and heavy smokers, light ex-smokers, and moderate and heavy ex-smokers. To determine the risk of death due to cigarette smoking, Roger et al matched data of a health survey to mortality data. Discrete time hazard models were employed to compute the risks. To examine the relationship between smoking and mortality the authors matched data of a health interview survey to mortality data.

The next step is to determine how many people exist in each smoking status (\(n\)):

\[ n = p \times \text{Pop}, \text{being Pop the age-specific population in of the area studied.} \]

The last step is to estimate the excess risk of death \(\oplus( R )\) of each smoking status relative to never smokers:

\[ m_{x,c} - m_{x,n}, \text{where } m_{x,c} \text{ is the age-specific central death rate for each smoking status and } m_{x,n} \text{ is the age-specific central death rate relative to never smokers.} \]

Finally the excess number of deaths is calculated as follows:
ED = \sum n^i (m_{x,c} - m_{x,n}), in the different ages-groups considered.

This method has been used once to estimate tobacco attributable mortality (88).
Predictive models

Under this head, one method was in evidence, namely These models are represented essentially by one model: the Prevent model (83)(88). Owing to the reduced scale on which this has been used or their use for more specific cases, their calculation process is not as well-known.

The Prevent simulation model (83)(88) was developed in 1988 in The Netherlands and is regarded as being a multifactorial generalization of the etiologic fraction. It has basically been used basically to predict mortality due to various causes, including tobacco (89)(94). The methodology used allows, among other factors, for a temporal dimension to be considered and envisages the possibility of a risk factor being-to associated with more than one disease and a disease being-to associated with more than one risk factor, i.e., multiplicity of cause or effect. The process of calculation is tedious and calls for knowledge of multiple data, such as birth- and mortality-rate series or the likelihood of dying at different ages for each sex (90)(95). [Suggested by Rev 1, Rev 4] The calculation procedure was described in detail in a PhD dissertation (90) and is summarized elsewhere (83, 91). Due its scarce use, the calculation procedure is not described in this paper. However, it is important to introduce two epidemiological effect measures that this method uses: The estimations are based on the “potential impact fraction” and in the “trend impact fraction”. Both are indicators of the reduction in the incidence of a disease in the study population, the former reflects changes in the evolution of a disease after an intervention and the latter is referred to autonomous or natural These trends. indicators are measures similar to the population attributable fractions, but in these cases zero is not the reference exposure category.
DISCUSSION

This paper constitutes our knowledge this paper represents of the first methodological review of procedures for estimating smoking-attributable mortality. In the context of decision-making, it is essential to know, albeit approximately, the impact that a given risk factor has on the mortality of a population. Estimation of tobacco-related mortality is not confined to one procedure alone, inasmuch as any of the different methods outlined above can be used for this purpose.

Despite the fact that different varying methodologies have been found, the foundations of more of them are the same and only a few differences arise in the calculation procedures. Data availability has been taken into account when choosing a method, as have and also methodology limitations and assumptions have to be considered. Some of them four general limitations are reviewed described below:

The first limitation affecting intercomparison of methods and studies stems from the absence of a universal definition of the categorization of tobacco use. The publications analyzed furnish different definitions of “smoker”, “non-smoker” and “ex-smoker” (6, 92-95), something that inevitably determines the result of the estimation (96).

To view smokers as a single entity would lead to a distorted mortality estimate, since failure to take account [Rev 4] of the dose-response gradient may over- or underestimate the mortality figures so attributed of the number of cigarettes smoked, age at initiation, years of smoking and other variables that could modify risks values can occur. It would thus be interesting to further explore tobacco use in the studied populations prior to the start of the estimation (88). Likewise, at a correct
The classification of ex-smokers is very important for estimating and predicting mortality attributable to tobacco use. To avoid overestimation of attributed mortality, Anthonisen (97)(100) proposed the account must be taken of one must consider the decrease in risk that takes place at 15 years after quitting the habit. Worth noting, however, is that this decrease is also determined by the subject’s age at cessation (98, 99)(101), the duration of smoking (100)(102) and the cause studied. The fact that this information was not expressly gathered in the majority of surveys means that mortality among ex-smokers may be overestimated. This problem is resolved, at least in part, by ex-smokers reclassifying themselves as non-smokers after the elapse of a certain minimum period of time without smoking (101)(103).

The second limitation, present mainly lies in the proportional method, resides in their reliance on current smoking prevalences to reflect mortality occasioned by tobacco use in previous years. Knowing current smoking prevalence could be a great help when it comes to predicting future mortality, but not present mortality (102)(104); indeed, knowing the prevalence of tobacco use in any given year could help predict lung cancer mortality in 20 years’ time (103)(105). As of yet, this problem has no easy solution, due to the absence, in most countries, of historical series of smoking prevalence in most countries. Moreover, even if such series were to exist, lack of knowledge of the latency and induction times for each of the tobacco-related causes of death would constitute another problem. The use of current prevalence may overestimate or underestimate the attributable mortality. In countries where the prevalence is decreasing, such as the U.S.A. or some European nations, the use of current prevalence is conservative, in the proportional attribution method, is conservative. The opposite occurs in countries where prevalence is increasing. Given the unavailability or inaccuracy of prevalence
data, and that emphasizing current prevalence is a poor proxy for cumulative hazards of smoking. Peto’s method is a good alternative in mortality estimation, the knowledge of the period of time from tobacco consumption until mortality related to tobacco consumption it is necessary.

Ascertaining the induction period might be feasible if only one specific component cause was active in triggering the disease. However, if one allows for the presence of more than one component cause, then each may have its own induction time. Furthermore, the action of effect modifiers variables could alter the duration of induction period. It would therefore seem that ascertainment of the induction period is complicated; and nevertheless, ascertainment of the latency period is no small easy matter either, since it varies according to the diagnostic methods employed.

What should be clear, however, is that an induction time is needed for tobacco to cause harm, and it is for this reason that the age ranges between 30-35 years are considered as one of the ages from which the appropriate time to begin measuring the effects of exposure. Measuring such effects without taking into account an induction time could lead to overestimated mortality results. On the other hand, some authors feel that ignoring mortality under the age of 35 years may give rise to underestimates of mortality figures, due to the existence of individuals who started smoking at early ages.

Peto and colleagues avoided the problem entailed in prevalence-dependent methods of attribution. For the application of their estimation procedure, lack of knowledge of the tobacco consumption or latency and induction periods are no a limitation. But this method has not been exempt from criticism, directed,
mainly, at the calculation of summarized prevalence. These authors Peto and colleagues define summarized synthetic prevalence as an indicator that summarizes a population’s smoking history, and calculated it by assuming CPS II data on lung cancer mortality rates among smokers and non-smokers to be valid. The use of these 2 sets of data gave rise to numerous criticisms that highlighted the low population representativeness of the CPS II (25, 107, 112, 113)(25, 109, 114, 115). Most of the population included in this cohort study was categorized as middle class, which may result in lung cancer mortality in non-smokers being underestimated (88)(93) leading, in turn, to an overestimation of lung cancer mortality attributable to tobacco use and, by extension, to an overestimation of the summarized prevalence (25). [Rev 4] Some of these critics were supported by the tobacco industry. To justify their validity and universality, these data were compared with those yielded by the study that targeted British physicians (93)(96). Despite the fact that the results obtained were similar, no conclusion could be drawn, since the representativeness of this latter study was also limited. The only thing that could be conclusively stated was that the lung cancer mortality rate among non-smokers had not varied over the years (112)(114). Nonetheless, in countries where the use of coal is widespread, lung cancer mortality among non-smokers is higher, and thus the data, rather than being drawn from the CPS II, have been drawn from a local study (72)(76).

The third limitation centers on the absence of world-wide risk indicators that would reflect the degree of association between tobacco and smoking related-causes of mortality. The most widely used effect measure is RR, and a sensitivity analysis has shown that changes in its value lead to a greater impact on the estimation of mortality than do changes in prevalence (102)(404). Although drawn from different sources, the
RRs used in the various studies mainly came from the CPS II (55, 56). Applying these risks to populations other than that of the USA aroused criticism because, inter alia, of their only being adjusted for age and sex, and because of the difficulty inherent in assimilating identical tobacco consumption and genetic variability patterns, or the same influence of confounding factors or effect modifiers. A solution to these problems was sought through a re-analysis of the data (9, 114-116), and the RRs were shown declared to be robust. Notwithstanding this, the criticisms continued unabated (117).

The risks obtained from the CPS II are plausible in the light of current knowledge (25) and have been extrapolated (118) to different EU countries, in absence of other high quality indicators. Nevertheless, other authors have chosen to apply RRs which are drawn from studies with less robust not appropriated designs (121), or possibly inconsistent with present knowledge (114, 117).

A fourth limitation of the methodology of attribution methodology is the uncertainty present in the relationship between exposure, tobacco use, and different causes of death. While lung cancer was the first disease to be causally associated with tobacco use, many studies have observed more causal associations. The latest report of the Surgeon General (56) has added 2 further causes of mortality that had not been considered to date, i.e., stomach cancer and acute myeloid leukemia and excludes hypertension.

Some methods have been compared by applying them in the same population. Published comparisons are the direct method individual analysis and SAMM Ec (3, 4), Peto and Prevent methods (83), Peto and proportional attribution method (11), and Garfinkel’s and proportional attribution method (85). The results obtained in all of
these comparisons have proved to be similar estimations, thereby conferring validity on the respective methodologies. Observational epidemiology and, despite their limitations, the use of the above-described calculation procedures offer a good approximation of the impact of tobacco on the mortality of a population (4).

[Highlighted] Until recently, no account was taken of the importance of genetics in the establishment of causal links. Methodologies such as Mendelian randomization provide—far better than could observational epidemiology—a robust approximation for understanding the effect of modifiable exposures on health (122). Identification of different genes that might be related to the higher or lower susceptibility to tobacco use would amount to an advance in the attribution of mortality.
CONCLUSIONS

Prior to conducting a study on estimation of tobacco-attributable mortality, it is essential to assess which method is best suited to the type and quality of the available information.

When the mortality estimation objective is going to be the knowledge of tobacco impact on a population, it is important to take into account all the diseases related with consumption. For this reason, the applications of methodologies that involve all the causes of disease are important. In sum, these methodologies are: Proportional attribution, Prevalence-based analysis in cohorts and in case-control studies, Peto and colleagues’, Lam’s and Roger’s methodology. All of them supply accurate and reliable estimations of mortality attributed to tobacco consumption.

The absence of a simulation study involving and comparing all calculations procedures does not allow us to recommend one a method over another one.

In all likelihood, the exact number of deaths occasioned by tobacco use will never be known, but these types of methods furnish estimates that constitute valuable information and help forming a more accurate picture of the problem that smoking poses to world health.
ACKNOWLEDGMENTS

The study was undertaken with the aid of a grant from the Galician Directorate-General for Public Health (Dirección Xeral de Saúde Pública).

We should like to thank Dr Alberto Malvar for his unstinting help throughout the course of drafting this manuscript, in the form of contributing ideas as well as reading the different versions produced. In addition, sincere thanks must go: to Prof Alberto Ruano-Ravina for his critical review of the manuscript and invaluable comments; and to Soly Santiago and Xurxo Hervada for their painstaking proof reading and suggestions.
REFERENCES


Prevalence of smoking is a poor proxy for cumulative hazards of smoking. Defining SIR (Smoking impact ratio or Synthetic prevalence) authors avoid prevalence limitations. By designing a case-control study OR could be assessed. RR extrapolation to populations different than studied is inconsistent. A calculation procedure let us to calculate the RR:

**Calculation procedure**

\[
SIR = \frac{C_{LC} - N_{LC}}{S_{LC}^* - N_{LC}^*}
\]

where \(C_{LC}, N_{LC}, S_{LC}^*, N_{LC}^*\) are age-sex specific lung cancer mortality rates for smokers and never smokers in the study and in the reference population (*).

\[
OR = \frac{a_i b_0}{a_0 b_i}
\]

\[
RR = \frac{a_i (1 - p_i)}{a_0 p_1}
\]

where \(p_i\) is the prevalence between the cases and:

<table>
<thead>
<tr>
<th></th>
<th>Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposed</td>
<td>(a_i)</td>
<td>(b_i)</td>
</tr>
<tr>
<td>Non exposed</td>
<td>(a_0)</td>
<td>(b_0)</td>
</tr>
</tbody>
</table>

**Packs-function in smokers**

\[
RR_s = 1 + ac((t-5) - t_0)
\]

**Multistage-function in smokers**

\[
RR_s = 1 + (t-5)^{4.5} + ac(1 + 2ac)((t-5) - t_0)^{4.5} + 2ac((t-5) - t_0)^{4.5})/(t-5)
\]

Where \(a\) is a constant, \(c\) is the number of packs of cigarettes smoked per year, \(t\) is the current age and \(t_0\) is age at start of smoking.

In former smokers \(t_1\) replaces \((t-5)\) and \(t_0\) is age at stop smoking.

Table 2.- Methodologies’ modifications taking into account proportional attribution method as base method.
<table>
<thead>
<tr>
<th>Method</th>
<th>Data employed</th>
<th>Data source</th>
<th>Method applied to estimate mortality due to:</th>
<th>Weaknesses</th>
<th>Strengths</th>
<th>Estimations calculated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportional attribution / SAMMEEC (n=52)</td>
<td>Prevalence</td>
<td>National Statistics</td>
<td>Tobacco consumption, exposure to environmental tobacco smoke (ETS), obesity, alcohol intake,…</td>
<td>- Does not take latency into account. - Generalizes the RR obtained in the Cancer Prevention Study (CPS)</td>
<td>- Worldwide use. - Comparability</td>
<td>Attributable mortality for all causes.</td>
</tr>
<tr>
<td>Prevent method (n=2)</td>
<td>Composition of the population</td>
<td>National Statistics</td>
<td>Tobacco consumption and general scenarios of effective health promotion.</td>
<td>- High need of information.</td>
<td>- Takes into account the multiplicity of cause or effect. - Proportional decrease in risk reduction related to time. - To measure the results of intervention policies.</td>
<td>Attributable mortality for all causes.</td>
</tr>
<tr>
<td>Hybrid method by Lam (n=4)</td>
<td>Mortality observed</td>
<td>National Statistics</td>
<td>Tobacco consumption and exposure to ETS.</td>
<td>- Case-control study design. - Recall bias.</td>
<td>- Specific risk dates.</td>
<td>Attributable mortality for all causes.</td>
</tr>
<tr>
<td>Garfinkel´s method (n=2)</td>
<td>Mortality observed</td>
<td>National Statistics</td>
<td>Tobacco consumption and alcohol intake.</td>
<td>- Partial view of the attributable mortality (only used to estimate mortality by cancer). - Assumes constant worldwide cancer mortality rates among never smokers.</td>
<td>- Necessary dates are few. - Low information requirements.</td>
<td>Cancer deaths attributable to smoking.</td>
</tr>
<tr>
<td>Rogers’ method (n=1)</td>
<td>Mortality observed (all causes)</td>
<td>National Statistics</td>
<td>Tobacco consumption.</td>
<td>- Availability of mortality registers. - Has a population representative survey about health-risks. - Assumption: smoking status remains steady since the survey about health-risks.</td>
<td>- Risks calculated ad hoc. - The population division is more reliable.</td>
<td>Attributable mortality for all causes.</td>
</tr>
</tbody>
</table>

Table 1.- Methods used to estimate tobacco attributable mortality.