Measuring the accumulated hazards of smoking: global and regional estimates for 2000

M Ezzati, A D Lopez

Objective: Current prevalence of smoking, even where data are available, is a poor proxy for cumulative hazards of smoking, which depend on several factors including the age at which smoking began, duration of smoking, number of cigarettes smoked per day, degree of inhalation, and cigarette characteristics such as tar and nicotine content or filter type.

Methods: We extended the Peto-Lopez smoking impact ratio method to estimate accumulated hazards of smoking for different regions of the world. Lung cancer mortality data were obtained from the Global Burden of Disease mortality database. The American Cancer Society Cancer Prevention Study, phase II (CPS-II) with follow up for the years 1982 to 1988 was the reference population. For the global application of the method, never-smoker lung cancer mortality rates were chosen based on the estimated use of coal for household energy in each region.

Results: Men in industrialised countries of Europe, North America, and the Western Pacific had the largest accumulated hazards of smoking. Young and middle age males in many regions of the developing world also had large smoking risks. The accumulated hazards of smoking for women were highest in North America followed by Europe.

Conclusions: In the absence of detailed data on smoking prevalence and history, lung cancer mortality provides a robust indicator of the accumulated hazards of smoking. These hazards in developing countries are currently more concentrated among young and middle aged males.

Smoking has been causally associated with substantially increased risk of premature mortality from lung cancer as well as other medical causes. Many of the health effects of smoking depend on the exposure history including the age at which smoking began, the number of cigarettes smoked per day, the degree of inhalation, and cigarette characteristics such as tar and nicotine content or filter type. Therefore, current smoking prevalence alone is an insufficient indicator of accumulated risk from smoking, even if data were available in all countries. Although such pattern variables have been studied in a few industrialised countries' few data are available elsewhere. Coupled with the fact that smoking has recently increased in many developing countries, this motivates using an exposure variable for global estimates of the burden of disease caused by smoking that better describes accumulated risk of smoking in countries with varying smoking histories.

Peto et al used absolute lung cancer mortality—for which data are available—to obtain the proportions of mortality from lung cancer as well as various other diseases attributable to smoking. We extended the Peto-Lopez method for indirect estimation of accumulated hazard caused by smoking to all regions of the world and used the smoking impact ratio (SIR) as a marker for accumulated smoking risk. This provides a comprehensive update of the estimates which were the basis for the Global Burden of Disease (GBD) 1990 tobacco mortality, using considerably more data on lung cancer mortality from GBD 2000 as well as new evidence on background lung cancer mortality from the Chinese retrospective proportional mortality study.

METHODS AND DATA SOURCES

Lung cancer mortality as indicator of accumulated smoking

Peto et al observed that the level of lung cancer mortality compared with never-smokers is an indicator of the "maturity" of the smoking epidemic in a population. The relation between cumulative smoking and lung cancer from various populations confirms this relationship. Based on this observation, SIR is defined as population lung cancer mortality in excess of that of never-smokers, relative to excess lung cancer mortality for a known reference group of smokers. Formally, the ratio in equation 1 below measures the absolute "maturity" of lung cancer mortality caused by smoking (that is, excess lung cancer mortality) in the study population, relative to the absolute "maturity" of lung cancer mortality in lifelong smokers of the reference population.

\[
SIR = \frac{C_{LC} - N_{LC}}{S_{LC} - N_{LC}} \quad (1)
\]

\(C_{LC}\): (age-sex-specific) lung cancer mortality rate in the study population (for example, country of analysis)

\(N_{LC}\): lung cancer mortality rate for never-smokers in the same population

\(S_{LC}\) and \(N_{LC}\): lung cancer mortality rates for smokers and never-smokers, respectively, in a reference population.

Liu et al found that in China the relative risk of mortality from lung cancer as a result of smoking is approximately constant in different cities whose non-smoker lung cancer mortality rates varied by a factor of 10 (see fig 4 in Liu et al). A constant relative risk over a large range of background mortality rates means that smoking results in a larger absolute excess mortality (that is, the numerator of equation 1) where

**Abbreviations:** CPS-II, American Cancer Society Cancer Prevention Study, phase II; ETS, environmental tobacco smoke; GBD, Global Burden of Disease; SIR, smoking impact ratio; WHO, World Health Organization
never-smoker lung cancer mortality is higher (and smaller absolute excess mortality where never-smoker lung cancer mortality is lower). Therefore, to be converted to an indicator of the maturity of the smoking epidemic, the numerator and denominator of equation 1 need to be normalised with the respective never-smoker lung cancer mortality rates. We define the background adjusted SIR by the following relationship:

$$SIR = \frac{G_{IC} - N_{IC}}{S_{IC} - N_{IC}} \times \frac{N_{IC}}{N_{IC}}$$  \hspace{1cm} (2)

where $c_{IC}$, $n_{IC}$, $s_{IC}$ and $n_{IC}$ are defined as above.

Following Peto et al.,\textsuperscript{14} we used the American Cancer Society Cancer Prevention Study phase II (CPS-II) population as the reference population. This is because among the numerous studies of smoking and cause-specific mortality, CPS-II is one of the very few conducted when the smoking epidemic was at its highest levels, especially for men. Therefore, the vast majority of (male) CPS-II current-smokers had been lifelong cigarette smokers. Further, the estimates of increased risk of mortality among smokers are available for both men and women and in smaller age groups than in other studies of mortality among smokers. Further, the estimates of increased risk of mortality among smokers are available for both men and women and in smaller age groups than in other studies of mortality among smokers.

It is straightforward to show that $SIR$ equals the proportion of the reference population (that is, CPS-II) smokers in a mix of smokers and never-smokers which has the same lung cancer mortality rate as the study population.\textsuperscript{15} This provides a convenient interpretation of $SIR$: using excess lung cancer mortality of a population over never-smokers, $SIR$ captures the accumulated hazards of smoking by converting the smokers in the study population into equivalents of smokers in the reference population where hazards for other diseases have been measured.\textsuperscript{16}

### Population and mortality statistics

The age-sex-specific population estimates for the 191 World Health Organization (WHO) member states were obtained from the United Nations Population Division and mortality statistics from WHO's GBD database. GBD mortality statistics are divided by country, sex, age (five year age groups up to 85 and 85+), and more than 150 causes of death. The 191 WHO member states are divided into 14 subregions based on a combination of WHO regions and epidemiological characteristics (child and adult mortality) (table 1).

The sources of mortality data for each region include vital registration and sample registration, incidence registries, and epidemiological studies. Details of mortality and cause-of-death analysis methods are provided elsewhere,\textsuperscript{17} and are summarised below as relevant. Mortality statistics by cause, region, and broad age groups are provided in the annexes of the World health report adjusted on an annual basis. The reliability of $SIR$ is determined by the reliability of lung cancer mortality estimates. In countries with good vital registration and medical certification of deaths (approximately 70 countries), lung cancer mortality is diagnosed with a high degree of accuracy. For example, microscopic confirmation of diagnosis against the cause reported on death certificates has suggested a 95% or higher confirmation rate in these settings.\textsuperscript{18} In approximately another 40 countries, vital registration of mortality was incomplete and medical certification of the cause was much less reliable. Standard demographic techniques are used in the GBD project to correct total death rates by age for these populations, and lung cancer rates are adjusted on an annual basis. In the absence of evidence for bias in mortality recording, the same correction factor is applied for all causes of death. Finally, for countries

<table>
<thead>
<tr>
<th>Table 1 Global Burden of Disease 2000 subregions. Source: World Health Organization (WHO)\textsuperscript{43}</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>WHO region</strong></td>
</tr>
<tr>
<td>AFR</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>AMR</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>EMR</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>EUR</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>SEAR</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>WPR</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

* A, very low child mortality and very low adult mortality; B, low child mortality and low adult mortality; C, low child mortality and high adult mortality; D, high child mortality and high adult mortality; E, high child mortality and very high adult mortality.
Estimating accumulated hazards of smoking

without vital registration, overall age-specific death rates are first determined using model life tables. Total cancer death rates are then estimated based on regional information about proportionate cancer mortality. Within this death rate, the distribution by site is based on regional incidence patterns from cancer registries reporting to the International Agency for Research on Cancer. Clearly, this indirect procedure has considerable uncertainty. Nevertheless, it is preferable to estimating mortality directly from reports of deaths to cancer registries which would very likely substantially underestimate mortality from lung and other cancers. Further, in the majority of the populations to which these indirect procedure is applied, lung cancer mortality is still low, hence constraining the estimate.

American Cancer Society CPS-II

The American Cancer Society’s CPS-II is a prospective study of smoking and death in more than one million Americans aged 30 years and older when they completed a questionnaire in 1982, with the latest follow up in 1998. A complete description of the study is provided elsewhere. In 1992, when the first six years (1982 to 1988) results were obtained, mortality follow up was virtually complete for the first two years, and about 98–99% complete for the next four. Because some conditions that cause death in the first two years may have affected smoking habits at entry (for example, those diagnosed with lung cancer may have stopped smoking because of their disease or related symptoms), analysis was restricted to years 3–6 inclusive (1984–88). The analysis related deaths (subdivided by cause, sex, and five-year age groups at the time of death) to person-years (with accounting for incompleteness) for those who in 1982 had never smoked regularly, and for those who were then current cigarette smokers. Most of the CPS-II current-smokers were lifelong cigarette smokers with a mean consumption of about 20 cigarettes per day.

SIR estimation

SIR values were calculated for individual countries and then averaged (population weighted) for each age group and sex in the 14 GBD subregions. The age groups used in the analysis were 0–4, 5–14, 15–29, 30–44, 45–59, 60–69, 70–79, and 80+ years. No effects before the age of 30 were attributed to smoking. The age groups used in the analysis averaged (population weighted) for each age group and sex in SIR estimation mean consumption of about 20 cigarettes per day. CPS-II current-smokers were lifelong cigarette smokers with a high likelihood of never-smoker lung cancer mortality in the study population.

Choice of never-smoker lung cancer mortality in the study population

In equation 2, $C_{ly}$ was obtained from the GBD mortality database and $S_{ly}$, $N_{ly}$, and $N_{ly}$, directly from CPS-II (see annex table in Peto et al). The only parameter to be estimated indirectly was $N_{ly}$, since direct estimates of never-smoker lung cancer mortality were available for very few countries. Figure 1 shows non-smoker lung cancer mortality for the USA (from CPS-II) and China. The detailed data from the Chinese retrospective proportional mortality study further allow dividing the Chinese rates into urban and rural areas, and the latter into coastal and inland.

As seen in fig 1, age-specific non-smoker lung cancer mortality was considerably larger in China than in the USA for both males and females. In China itself, there were notable differences between different parts of the country, with urban areas having the highest non-smoker lung cancer mortality rates, followed by the rural areas on the coast and inland rural areas (see also fig 4 in Liu et al). The variability of the non-smoker lung cancer mortality rates, between the USA and China as well as between urban and rural regions of China is explained by the Chinese patterns of household energy use over the past few decades. Coal is a common household fuel in China, often burned in stoves and buildings without adequate ventilation. Exposure to coal smoke and cooking fumes has been associated with increased lung cancer incidence in China.

In inland rural regions of China, where incomes are lowest, biomass (for example, crop residues and wood) has been the dominant household fuel, while in the coastal villages and cities the use of coal has been more common.

The relationship between wood smoke and lung cancer has been absent or considerably weaker than that between coal and lung cancer in epidemiological studies. Although urban air pollution has been linked to increased lung cancer mortality in some studies, the size of the risk has been considerably smaller than the effects of smoking or direct exposure to coal smoke. Further, the risk from urban air pollution on lung cancer has been found to be smaller among non-smokers than among smokers, with one study finding increased risk among smokers only. For example, exposure to high levels...
of urban ambient air pollution in Poland, even in regions
where coal is used extensively, increases the risk of mortality
from lung cancer by approximately 14% among non-smokers
(versus 40% among smokers). Coupled with the fact that
only small fractions of national populations live in the most
polluted urban areas, the lung cancer mortality of non-
smokers is not expected to be greatly affected by ambient air
pollution. For example, the lung cancer mortality rate among
non-smoking American women remained constant at ap-
proximately 12 per 100,000 between 1960 and 1986, despite
the preceding changes in exposure to urban air pollution.

Based on the pattern of background lung cancer mortality
rates, and the underlying mechanisms of increased lung can-
cer mortality in China and its various regions, background
(never-smoker) lung cancer rates for the different subregions
were based on the estimated use of coal for domestic energy in
vented stoves. We used Chinese non-smoker rates (fig 1)
for China, a weighted average of Chinese and CPS-II

Figure 2 Estimates of smoking impact ratio (SIR) by age, sex, and GBD subregions. Note that the vertical axis scales are different in the
different panels to increase resolution. The numbers next to the subregions in the legend indicate estimates of the regional prevalence. See
table 1 for key to subregions.
non-smoker rates for the Indian subcontinent where coal is also used for household fuel (with weights for Chinese rates equal to the prevalence of coal use), and CPS-II non-smoker rates for the remaining countries of the world where domestic coal use in unvented stoves is absent or negligible. The remaining risk factors with potential effects on lung cancer mortality (urban air pollution, occupational hazards, radon, wood smoke, etc) affect most populations to varying degrees. Therefore, the net impacts of these other risk factors are considered as sources of uncertainty in extrapolating non-smoker lung cancer mortality from one setting to another, rather than in determining mean never-smoker lung cancer rates.

RESULTS

Figure 2 shows SIR for females and males in developing and industrialised countries. WPR-B was further divided into China and all other countries in the region. Given that the population of the subregion is mostly from China, the regional estimates are dominated by the Chinese data. This division allows the characteristics of the other countries to also be considered. We have also shown in the figure legend an estimate of adult smoking prevalence for each region. Prevalence estimates are uncertain and based on data from a limited number of countries. Despite uncertainties in country level and age-specific estimates, regional estimates have been found to be relatively robust.16

A number of important features of global smoking epidemic are seen in the different panels of fig 2. The largest accumulated risk of smoking for males below the age of 70 was in the countries of Eastern Europe and the former Soviet Union (EUR-B and EUR-C). For older males (age ≥ 70 years), North American and Western European countries had the largest accumulated risk. These two results are consistent with the very high current and recent prevalence of smoking among Eastern European men, and the longer history of smoking among North American and Western European men.

The accumulated hazards of smoking among men were lowest in AMR-D, AFR-D, and AFR-E. In these regions, widespread smoking is a relatively more recent phenomenon with total consumption still low. The results for China are particularly important and instructive. The background adjusted SIR for China was fairly low, despite high lung cancer mortality in this country. This is consistent with the finding of the Chinese retrospective proportional mortality study4 that in 1990 the relative risk of lung cancer mortality for a Chinese smoker was less than 3.0 (compared to more than 20 for CPS-II smokers) because of the more recent tobacco epidemic in this country. Together with the finding of Liu et al3 that smoking acts to “amplify” the high background rates of lung cancer, the expected increase in accumulated hazard is likely to result in enormous lung cancer (and other) mortality in China. Note that if we had not accounted for the Chinese specific never-smoker rates, the SIR for men would be similar to those in SEAR-B and for women higher than any other developing country.

The comparison of SIR and prevalence also emphasises the importance of considering accumulated risks. For example, the prevalence of smoking among adult men in AMR-A was equal to AFR-D and lower than all other regions in the developing world. The SIR values for AMR-A males were, however, larger than those of most developing country regions because of the different histories of smoking. In AMR-A, smoking has been declining and current prevalence would underestimate the accumulated hazards of smoking. In developing countries, on the other hand, smoking has been rising in recent decades with current prevalence being comparable to AMR-A, but the accumulated hazards are still lower.

Among women, North America (AMR-A) had the single highest accumulated risk from smoking. Although women in some countries in Western Europe (for example, UK) have smoked for several decades, smoking is more recent in the southern parts of the continent. Therefore, the regional SIR is still lower than in AMR-A. SIR values for women were consistently low in developing countries except for younger and middle aged women in AMR-B (Latin America and the Caribbean) and young women in EMR-B.

Age patterns of SIR estimates for the different regions also provide information about the state of the smoking epidemic. Among men in industrialised countries, SIR was relatively constant across ages in AMR-A, and had a declining age trend in other industrialised regions, with EUR-A being closest to the constant pattern. These age patterns within and between regions mean that, relative to the same age group of American smokers in the 1980s, hazards of smoking have been relatively constant in North America across age groups whereas smoking has been more concentrated among younger and middle aged men in Europe and the Western Pacific, especially in the former Soviet Union (EUR-C). In the developing countries of Latin America and the Caribbean, the Eastern Mediterranean, and sub-Saharan Africa, smoking hazards had a fairly constant age profile, each compared to the same age group of American smokers in the 1980s. In Asia, on the other hand, smoking exposure among the younger and middle aged male cohorts was greater than the older groups, when compared to the same age group of American smokers in the 1980s.

The age patterns of SIR were overall more constant among women in both developing and industrialised countries compared to men. In AMR-B, EMR-B, EMR-D, and the European regions SIR was higher among younger and middle aged women than the older ones, each compared to the same age group of female American smokers in the 1980s. In Asia, female SIR peaked among the middle aged cohorts, which may reflect social factors that would prevent smoking among many young women. In North America, the distribution was more towards older age groups when compared to the same age group of female American smokers in the 1980s, indicating reduced smoking in younger ages relative to the older group.

DISCUSSION

We extended the indirect Peto-Lopez method,17 which uses absolute lung cancer mortality in a population as a marker for accumulated hazards of smoking, to estimate the accumulated risks of smoking in different regions of the world. We chose the never-smoker lung cancer mortality based on direct estimates or by extrapolation from other regions based on best available evidence, explicitly stating the assumptions and reasons for each choice.

Sources of uncertainty in SIR estimates include lung cancer mortality data from the GBD database, never-smoker lung cancer mortality estimates, and reference population (CPS-II) smoker and never-smoker lung cancer mortality. More broadly, cigarette smoke characteristics that cause lung cancer may not be fully correlated with those that cause other hazards (for example, carcinogens in cigarette smoke are the cause of cancer but the concentration or size distribution of respirable particles may be a better indicator of other health effects). The validity of the SIR method has nonetheless been confirmed against other methods.18,19,20 Similarly, the delay in hazard may be different for lung cancer and the other diseases caused by smoking. Lung cancer risk rises over an extended period of smoking. This may be shared by some other outcomes of smoking (for example, other cancers and chronic obstructive pulmonary disease) but other diseases may occur after immediate exposure (for example, acute respiratory diseases) or shorter delays (for example, vascular disease). When smoking is on the rise, an indicator based on excess lung cancer would underestimate the impacts of those diseases that occur earlier, and vice versa. Since smoking has been increasing in most regions of the world, the net effect of this will be
an underestimation of current mortality caused by vascular and other diseases when using SIR as the indicator of accumulated hazard. Exposure to environmental tobacco smoke (ETS) introduces additional uncertainty in SIR estimates as the measure of accumulated hazards of smoking. ETS exposure is a likely cause of lung cancer and other diseases. Of the four variables in the SIR relationship (\(C_t, N_{t-}, S_{t-}, N_{t} \), in equation 2) the effects of ETS are smallest on \(C_t \), which is almost completely affected by direct smoking. \(C_t \), captures the effects of both direct and indirect exposure to smoking but is affected more by direct smoking. The effects of ETS are largest on \(N_{t-}\), never-smokers (\(N_{t-}\)) who would otherwise not be exposed to tobacco smoke. Therefore, both \(N_{t-}\) and \(N_{t} \) are larger than they would have been in the absence of ETS. Since \(C_t \) is almost always smaller than \(S_{t-}\), increasing \(N_{t-}\) and \(N_{t} \) would result in a larger relative reduction (through subtraction) in the numerator of equation 1 than in the denominator and therefore underestimate SIR values.

Using this method, we estimated that in 2000 the relative accumulated hazards of smoking were highest among men in industrialised countries, in particular among the young and middle age (30–69 years) cohorts in Eastern Europe.* This age group also had relatively large accumulated risks from smoking in many developing countries, especially in Asia and Latin America, reflecting increases in smoking in these regions in recent decades. For females, the accumulated hazards were still low in most regions with the exception of North America and younger and middle age women in Europe and Latin America and the Caribbean.

Accumulated risks of smoking exhibited large variation among different geographical regions, in particular in the developing world. This inter-regional variation, which was larger than that observed in industrialised countries, occurs because the shape and maturity of the smoking epidemic has been highly affected by the varying economic and cultural determinants of smoking in these settings. A few general statements can nonetheless be made about the health impacts of smoking in these countries. First, relative to industrialised countries, developing countries had a higher proportion of smoking attributable risk in the 30–69 year age group than in the 70+ group. This shows that as people (mostly men) who began smoking over the past three decades in developing countries become older, mortality caused by smoking is likely to rise further. Second, the current risks caused by smoking in this setting were highly concentrated among men. Given the shifting disease patterns, and because most of the growth in global population is expected to take place in the developing world, the health loss from smoking, already one of the most important global health hazards, will continue to grow unless effective interventions and policies that curb and reduce smoking among males and prevent increases among females in these countries are implemented.

ACKNOWLEDGEMENTS
This work has been sponsored by the National Institute of Aging grant 1-PO1-AG17625. We thank J Boreham, J Henley, R Peto, and M Thun for discussions on methodology and for additional re-analysis of data from CPS-II and Chinese retrospective proportional mortality study, and E Guindon and P Jha for data on prevalence of smoking. This paper has benefited from comments by four anonymous reviewers.

Authors’ affiliations
†M Ezzati, Risk, Resource, and Environmental Management Division, Resources for the Future, Washington, DC, USA
‡Now at School of Population Health, University of Queensland, Brisbane, Australia

What this paper adds
This paper extends the Peto-Lopez method to estimate accumulated hazards of smoking for different regions of the world. The paper uses the comprehensive, and regularly updated, database of the Global Burden of Disease project for lung cancer mortality. Using a critical review of factors other than smoking that affect lung cancer, it also adjusts for these factors in the model based on household use of coal in unvented stoves. The paper provides a consistent source for estimating the accumulated hazards of smoking in industrialised and developing countries, including for monitoring the risks associated with smoking over time.

REFERENCES
16 Bennett NG, Horuchi S. Mortality estimation from registered deaths in less developed countries. Demography 1984; 21: 217–34.
Estimating accumulated hazards of smoking


