Smoking prevalence in WHO Member States

Monitoring the prevalence of tobacco use is central to any surveillance system for tobacco control. Reliable prevalence data provide the information needed to assess the impacts of tobacco control actions adopted by a country and can be used by tobacco control workers in their efforts to counter the tobacco epidemic. This report contains WHO-modelled prevalence estimates for smoking and country-provided data for both smoking and smokeless tobacco use (see Appendix VII).

Collection of tobacco use prevalence estimates

As discussed in Technical Note I, the data collection process for this report differed from that used in the previous two editions of the WHO Report on the Global Tobacco Epidemic. Information for the previous reports was collected by sending a detailed questionnaire to Member States about any tobacco survey that had been conducted in the relatively recent past. For this report, three sources of data were first explored: (i) reports submitted to the WHO FCTC Secretariat by Parties; and (ii) information sought from the WHO STEPwise Survey team (also responsible for the WHO Global Infobase, a portal of information on countries that may have conducted surveys and officially released the results of a recent survey implemented since 2005. In addition, an extensive literature search was conducted to identify any other possible data sources. During this process, multiple data sources were frequently identified. In such cases, preference was given to surveys that met the following four criteria:

- survey the adult population aged 15 years and older.
- Member States were contacted if WHO did not obtain an official report from the recently undertaken surveys.
- Data were collated on four indicators of tobacco smoking:
  - current and daily prevalence of tobacco smoking;
  - current and daily prevalence of cigarette smoking.
- These indicators provide the most complete representation of tobacco smoking across countries. Although differences exist in the types of tobacco products used in different countries and grown or manufactured in different regions of the world, data on cigarette smoking and tobacco smoking are the most widely available and are common to all countries, thereby permitting statistical analyses.

Analysis and presentation of tobacco use prevalence estimates

Data collected on prevalence estimates are presented in this report in two forms.

- Crude prevalence rates (Appendix VIII): these should be used to assess the actual use of tobacco in a country and to generate an estimate of the number of smokers for the relevant indicator (e.g. current smokers, daily smokers) in the population.
- Adjusted and age-standardized prevalence rates (Appendix VII): these rates are constructed solely for the purpose of comparing tobacco use prevalence estimates across multiple countries or across multiple time periods for the same country. These rates must not be used to estimate the number of smokers in the population. The methods for adjusting and age-standardizing for survey differences are described separately below, but the estimates presented in Appendix VII have been both adjusted and age-standardized.

Crude prevalence. The crude smoking prevalence, a summary measure of tobacco use in a population, reflects the actual use of tobacco in a country (e.g. prevalence of smoking by adults aged 15 years and older). The crude rate, expressed as a percentage of the total population, refers to the number of smokers per 100 population. When this crude prevalence rate is multiplied by the country’s population, the result is the number of smokers in the country.

Adjusted prevalence. Adjustments to data are typically done when collecting information from heterogeneous sources that originate from different surveys and do not employ standardized survey instruments. These differences render difficult the production of national-level age-standardized rates. WHO has also developed a regression method that attempts to adjust the estimates to enable comparisons of results between countries. The general principle that underlies the regression method is that if data are partly missing or are incomplete for a country, then the regression technique uses data available for the region in which the country is located to generate estimates for that country. The regression models are run at the United Nations subregional level separately for males and females in order to obtain age-specific prevalence rates for that region. These estimates are then substituted for the country falling within the subregion for the missing indicator. Note that the technique cannot be used for countries without any data; these countries are excluded from any analysis. The four types of differences between surveys and the relevant adjustment procedures used are listed below.

Differences in age groups covered by the survey. In order to estimate smoking prevalence rates for standard age ranges (by five-year groups from age 15 until age 80 and then aggregated from 80 to 100 years), the association between age and daily smoking is examined for males and females separately for each country using scatter plots. For this exercise, data from the
latest nationally representative survey are chosen; in some cases more than one survey is chosen if male and female prevalence rates stem from different surveys or if the additional survey supplements data for the extreme age intervals. To obtain age-specific prevalence rates for five-year age intervals, regression models using daily smoking prevalence estimates from a first-order, second-order and third-order function of age are graphed against the scatter plot and the best-fitting curve is chosen. For the remaining indicators, a combination of methods is applied: regression models are run at the subregional level to obtain age-specific rates for current and daily cigarette smoking, and an equivalence relationship is applied between smoking prevalence rates and cigarette smoking where cigarette smoking is dominant to obtain age-specific prevalence rates for current and daily cigarette smoking for the standard age intervals.

Differences in the types of indicators of tobacco use measured. If we have data for current tobacco smoking and current cigarette smoking, then definitional adjustments are made to account for the missing daily tobacco smoking and daily cigarette smoking data. Likewise, if we have data for current and daily tobacco smoking only, then tobacco type adjustments are made across tobacco types to generate estimates for current and daily cigarette smoking.

Differences in geographic coverage of the survey within the country. Adjustments are made to the data by observing the prevalence relationship between urban and rural areas in countries falling within the relevant subregion. Results from this urban-rural regression exercise are applied to countries to allow a scaling-up of prevalence to the national level. As an example, if a country has prevalence rates for daily smoking of tobacco in urban areas only, the regression results from the smoking relationship are used to obtain rural prevalence rates for daily smoking. These are then combined with urban prevalence rates using urban-rural population ratios as weights to generate a national prevalence estimate as well as national age-specific rates.

Differences in survey year. For this report, smoking prevalence estimates are generated for the year 2009. Smoking prevalence data are sourced from surveys conducted in countries in different years. In some cases, the latest available prevalence data came from surveys before the year 2009 while in other cases the survey was later than 2009. To obtain smoking prevalence estimates for 2009, trend information is used either to project into the future for countries with data older than 2009 or backtracked for countries with data later than 2009. This is achieved by incorporating trend information from all available surveys for each country. For countries without historical data, trend information from the respective subregion in which they fall is used.

In the absence of crude prevalence rates for the relevant indicator, adjusted prevalence estimates can be used to assess the number of smokers for the relevant indicator in a country.

Age-standardized prevalence. Tobacco use generally varies widely by sex and across age groups. Although the crude prevalence rate is reasonably easy to understand for a country at one point in time, comparing crude rates between two or more countries at one point in time, or of one country at different points in time, can be misleading if the two populations being compared have significantly different age distributions or differences in tobacco use by sex. The method of age standardization is commonly used to overcome this problem and allows for meaningful comparison of prevalence between countries. The method involves applying the age-specific rates by sex in each population to one standard population. When presenting age-standardized prevalence rates, both this and the previous WHO reports on the global tobacco epidemic used the WHO Standard Population, a fictitious population whose age distribution was artificially created and is largely reflective of the population age structure of low- and middle-income countries. The resulting age-standardized rate, also expressed as a percentage of the total population, refers to the number of smokers per 100 WHO Standard Population. As a result, the rate generated using this process is only a hypothetical number with no inherent meaning in itself. It is only useful when comparing rates obtained from one country with those obtained in another country, or from the same country at a different point in time. In order to produce an overall smoking prevalence rate for a country, the age-standardized prevalence rates for males and females must be combined to generate total prevalence. Since the WHO Standard Population is the same irrespective of sex, the age-standardized rates for males and females are combined using population weights for males and for females at the global level from United Nations population data for 2009. For example, if the age-standardized prevalence rate for tobacco smoking in adults is 60% for males and 30% for females, the combined prevalence rate for tobacco smoking in all adults is calculated as 60 x (0.51) + 30 x (0.49) = 45%, with the figures in brackets representing male and female population weights. Thus, of the total smoking prevalence (45%), the proportion of smoking attributable to males is 66.7% \(\frac{60}{100}\) and to females 33.3% \(\frac{30}{100}\). These combined rates are shown in Appendix VII.

1 Tobacco smoking includes cigarettes, cigars, pipes and any other form of smoked tobacco.
2 For countries where consumption of smokeless tobacco products is high, we have published these data for that particular country.
3 There are 21 United Nations subregions; Oceanía, Melanesia, Polynesia and Micronesia are combined into one subregion to form a total of 18. For a complete listing, please refer to World Population Prospects, 2008 revision at http://esa.un.org/unpp/index.asp?panel=5