Annex A: Data and methods

For *The world health report 2002*, WHO developed a new framework for quantifying deaths and burden of disease caused by risk factors, with an emphasis on improving the comparability of the estimates *(1)*. Different risk factors have very different epidemiological traditions, particularly with regard to defining "hazardous" exposure, the strength of evidence on causality, and the availability of epidemiological research on exposure and outcomes. Moreover, classical risk factor research has treated exposures as dichotomous, with individuals either exposed or not exposed, and with exposure defined according to a threshold value that is often arbitrary. Recent evidence for such continuous exposures as cholesterol, blood pressure and body mass index suggests that such arbitrarily defined thresholds are inappropriate because hazard functions for these risks change continuously across the entire range of measured exposure levels, with no obvious threshold *(e.g., 28)*.

The risk factor burden was calculated for *The world health report 2002* as the reduction in disease burden that would be expected under the risk factor exposure scenario that minimizes risk *(29, 30)*. Fractions of disease burden attributable to a risk factor were calculated based on a comparison of disease burden observed under the current distribution of exposure by age, sex and region, with that expected if a counterfactual distribution of exposure had applied. To improve comparability across risk factors, a counterfactual distribution was defined for each risk factor as the population distribution of exposure that would lead to the lowest levels of disease burden. This counterfactual exposure is assumed to have applied in the reference year and in all previous years.

For this update of global estimates of mortality and burden of disease attributable to 24 global risk factors, the methods developed for *The world health report 2002* were applied, with updated inputs on exposure distributions for 2004 and, in some cases, updated estimates of the magnitude of the hazards associated with specific risk exposures. These revisions are documented below. Two new risk factors were included for the first time in this update: suboptimal breastfeeding and higher-than-optimal blood glucose. Regional-level estimates of mortality and DALYs for specific diseases and injuries for 2004 were from a recent WHO update of global burden of disease estimates *(2)*. For some risk factors – including fruit and vegetable intake, occupation risk factors, child sexual abuse and unsafe health-care injections – revised estimates of exposure distributions were not available, and disease- and injury-specific population attributable fractions (PAFs) calculated for the year 2000 were assumed also to apply in 2004.

A1.1 Estimating population attributable fractions

To calculate the difference in population health under the counterfactual scenario, the PAF is first calculated. PAF is defined as:

$$PAF = \frac{\int_{x=0}^{m} RR(x)P(x)dx - \int_{x=0}^{m} RR(x)P'(x)dx}{\int_{x=0}^{m} RR(x)P(x)dx} \qquad (1)$$

where $RR(x)$ = relative risk at each exposure level, $P(x)$ = proportion of population at each exposure level, $P'(x)$ = counterfactual proportion of population at each exposure level, and m = maximum exposure level *(31)*.

For risk factors where the exposure is dichotomous (exposed, not exposed), and the counterfactual scenario is no exposure, equation 1 reduces to equation 2:

$$PAF = P \cdot (RR - 1) / [P \cdot (RR - 1) + 1] \qquad (2)$$

where P = prevalence of exposure, and RR = relative risk for exposed versus non-exposed.

Once the fraction of a disease (or injury) that is attributed to a risk factor has been established, the attributable mortality or burden is simply the product of the total death or DALY estimates for the disease and the attributable fraction. For most diseases, the same attributable fraction is applied to fatal (YLL) and non-fatal (YLD) burden estimates.

Choice of counterfactual

Analysis using counterfactual exposure distributions requires comparing the current distributions of exposure to risk factors with some alternative