Guidance note to facilitate country consultation on WHO/ILO Joint Estimates of the work-related burden of ischemic heart disease and stroke attributable to exposure to long working hours, for the years 2000, 2010 and 2016

March 2020

Environment, Climate Change and Health Department, WHO Headquarters

Background

The World Health Organization (WHO) and the International Labour Organization (ILO) have developed joint estimates of the work-related burden of disease and injury (WHO/ILO Joint Estimates). These comprise estimates of exposure to occupational risk factors and estimates of burden of disease attributable to these occupational exposures. To establish the evidence base for producing these estimates, WHO and ILO have conducted systematic reviews and compiled input data in databases.

Objectives

We aimed to estimate the proportion of the population exposed to (hazardous) long working hours and the burdens of ischaemic heart disease and stroke attributable to exposure to long working hours.

Objectives of the country consultation

This country consultation is a process through which WHO invites feedback from countries on its estimates. It is important that relevant authorities consult the data files to provide effective feedback.

Results

Feedback is invited on the following WHO/ILO Joint Estimates:

1. Proportion of the population exposed to (hazardous) long working hours;
2. Number of deaths due to ischemic heart disease attributable to exposure to long working hours;
3. Number of deaths due to stroke attributable to exposure to long working hours;
4. Number of disability-adjusted life years (DALYs) due to ischemic heart disease attributable to exposure to long working hours; and
5. Number of DALYs due to stroke attributable to exposure to long working hours.

These estimates are produced for three years (2000, 2010 and 2016) and disaggregated by long working hours category (2 categories: 49-54 and ≥55 hours/week); sex (3 categories: both female and male, female, and male); and age group (13 categories: ≥15, 15-19, … , 90-94, and ≥95 years).

Data sources

The estimates were produced using of the six sets of input data described below.

Input Data 1: Cross-sectional data on proportion of survey participants per working hours category

WHO and ILO specifically established the WHO/ILO Global Working Hours Database. Microdata were sourced on the total number of hours usually or actually worked per week in first and second (or all) jobs. These source data are 467 million observations from 2,324 surveys conducted in 154 countries between 1 January 1976 and 31 December 2018 (*Table 1*). Of all surveys, 96.4% are official surveys (primarily Labour Force Surveys) collected by national statistical offices, and 3.6% are Gallup surveys. All source data have been shared by countries with the ILO Department of Statistics; exceptions are that one official data set for China was shared by the National Institute of Occupational Health and Poison Control, Chinese Center for Disease Control and Prevention with the WHO Climate Change and

Health Department and the official data for Japan were extracted from an open-access database of the Statistics Bureau of Japan. Microdata on the number of working hours were extracted from the source surveys. They were harmonized into six working hours categories (*i*): 0 hours/week (labour market inactive) (coded as 0); 0-34 hours/week (1), 35-40 hours/week (2), 41-48 hours/week (3), 49-54 hours/week (4), and ≥55 hours/week (5)). They were weighted using the weights provided by countries and aggregated by population defined by country, year, sex and age group. The raw data were not modified in any way. The database is not publicly available to protect the confidentiality of the survey data it contains.

Table 1: Coverage of surveys and countries in the WHO/ILO Global Working Hours Database

	Region (defined as per WHO classification)						**World**
	AFRO	**AMRO**	**SEARO**	**EURO**	**EMRO**	**WPRO**	
Surveys (N)	135	437	96	1,435	66	155	2,324
Countries with ≥1 survey (N) (% of countries)	37 (78.7%)	24 (68.6%)	10 (90.9%)	45 (84.9%)	11 (50.0%)	27 (92.3%)	154 (77.4%)

Input Data 2: Longitudinal data on proportion of survey participants per working hours category and labour force activity category

WHO and ILO also specifically established the WHO/ILO Global Longitudinal Working Hours Database. Longitudinal microdata were sourced on the total number of weekly working hours. The database comprises 143 million observations from 739 waves of quarterly Labour Force Surveys conducted in 15 countries between 1 January 2000 and 31 December 2018 (Table 2). All these surveys were official surveys conducted by national statistical offices and have been shared by countries with the ILO Department of Statistics. These quarterly surveys use sample rotation to ensure sample overlaps, with measures taken repeatedly from the same survey participants over consecutive years. Because the microdata did not include individual participant identifiers, we probabilistically linked data longitudinally using matching by household number, household sequence number, sex and birth year. The microdata on the number of working hours were extracted; harmonized into the standard working hours categories; weighted using the weights provided by countries; and aggregated by population defined by country, year, sex and age group. Data on labour force activity were also extracted and included. Raw data were not modified, and to protect confidentiality the database was not published.

Table 2: Coverage of survey waves and countries in the WHO/ILO Longitudinal Global Working Hours Database

	Region						**World**
	AFRO	**AMRO**	**SEARO**	**EURO**	**EMRO**	**WPRO**	
Survey waves (N)	66	436	50	34	39	114	739
Countries with ≥1 survey wave (N) (% of countries)	2 (4.26%)	7 (20.00%)	2 (18.2%)	1 (1.98%)	1 (4.55%)	2 (6.9%)	15 (7.61%)

Input Data 3: Estimates of the total number of population

Estimates of the total number of population by country, year, sex and age group for the years 1950-2018 were sourced from the United Nations latest global population estimates (UN 2019).

Input Data 4: Estimates of probability of death

Estimates of probability of death by country, year, sex and age group were sourced from the WHO life tables (WHO 2020).

Input Data 5: Estimates of total number of deaths and disability-adjusted life years

Estimates of total number of deaths and disability-adjusted life years (DALYs) for ischemic heart disease and stroke for the years 2000, 2010 and 2016 were sourced from the WHO Global Health Estimates (WHO 2018).

Input Data 6: Estimates of relative risks

WHO and ILO, supported by a large network of individual experts, have conducted systematic reviews and meta-analyses of the relative risk of ischaemic heart disease (Li et al 2018, under review) and stroke (Descatha et al 2018, under review) among individuals exposed to long working hours, compared with those exposed to standard working hours of 35-40 hours/week (Table 3. These systematic reviews concluded that there is sufficient evidence that working 49-54 hours/week increases the risk of stroke and working ≥55 hours/week increases the risk of ischemic heart disease and stroke (Table 3).

Table 3: Relative risks of ischaemic heart disease and stroke by working hour category

Working hours category	Relative risk (95% CI)	
	Ischemic heart disease	**Stroke**
35-40 hours/week	Reference	Reference
41-48 hours/week	Insufficient evidence	Insufficient evidence
49-54 hours/week	Insufficient evidence	1.13 (1.00-1.28)
≥55 hours/week	1.17 (1.05-1.31)	1.35 (1.13-1.61)

Methods

The estimation strategy comprised modelling the input data (i.e., Input Data 1-6) in four distinct models that consecutively built on each other (Models 1-4).

Model 1: Multilevel model to estimate proportion of population in each working hour category at each year

For each year between 1980-2016, for each population defined by country, sex and age group, we produced estimates of the proportion of the population in each of the six standard working hour categories (*i*). We modelled Input Data 1 using the following multilevel model (Model 1):

$$proportion_i = A_i(sex, age, country) + B_i(sex, age, country) * year$$

where $proportion_i$ is the proportion of the population in the working hours category *i* in a given population defined by country, sex and age group; *year* is the survey year; and *sex*, *age* and *country* are the sex, age group and country of the population. The intercept $A_i(sex, age, country)$ and slope $B_i(sex, age, country)$ of the year dependence of $proportion_i$ are calculated with a multilevel model with sex and age as fixed effects and *sex* and *age* as random effects, nested in the *country* within the region (with regions treated independently). Because $proportion_i$ was strongly non-linearly dependent on *age*, we linearized *age* by 5th order orthogonal polynomials to prevent collinearity.

Model 2: Model of transition probabilities between working hours categories

For each population defined by country, sex and age group, we estimated the probability ($probability_j$) of transitioning from the working hours category *i* in $year_t$ to the working hour category *i* in $year_{t+1}$. The *j* denotes one of 36 possible transitions from one of the six working hour categories in $year_t$ to one specific working hour category in $year_{t+1}$.

We adopted the methods developed by Eurostat for calculating these transition probabilities (Eurostat 2020). Using Input Data 2, we scaled the survey weights for the target year ($year_{t+1}$) to represent the correct labour market status by country, sex and age group for the initial year (year$_t$) and the target year and then adjusted the complete sample in the target year to match margins for labour market status in both years, using iterative raking by sex. We did not match the working hours category *i* = 0 (labour market inactive) for the initial year. We modelled Input Data 2 using the following multinomial logit regression model (Model 2):

$$probability_j = \frac{e^{\beta_j X_j}}{1 + \sum_\alpha e^{\beta_\alpha X_j}}$$

where β_j is the set of regression coefficients describing the longitudinal weights associated with the transition *j*; X_j is a set of explanatory variables (sex and age as a fractional polynomial with maximal permitted degree of 4 associated with the transition *j*); and the summation (index α) goes through all possible transitions *j* (except the transition from *i* = 0 in year$_t$ to *i* = 0 in year$_{t+1}$ which was chosen as a pivot outcome).

With Model 2, we derived 15,900 transition probabilities for the 15 countries with data in Input Data 2. In addition, modelling quarterly European Union Labour Force Surveys using Model 2, Eurostat derived 31,104 transition probabilities covering 27 countries and shared these transition probabilities with WHO and ILO. For populations defined by country, sex and age group for whom $probability_i$ could not be calculated because the required longitudinal data were unavailable, $probability_i$ was imputed. The imputed $probability_i$ was the mean of all transition probabilities of the population defined by the same sex and age in the region, weighted by the number of observations contributing to the transition probabilities.

Model 3: Microsimulation model to estimate exposed population over time window

For each population defined by country, sex and age group, we estimated the proportion ($proportion_k$) of the population in each working hours category over the time window (*k*). We defined *k* as the highest working hour category *i* in any year in the time window. Based on advice from the WHO Technical Advisory Group, we assumed that burden of ischemic heart disease and stroke in the estimation year was the result of exposure to (hazardous) working hours over a time window of 15-5 years before the estimation year. For example, to estimate burden in the estimation year 2016, we assume that the time window of exposure was 2001-11.

We used microsimulation, a method for generating micro-level estimates by combining individual- and aggregate-level datasets. For each country, we initiated a synthetic population, using Input Data 3 to ensure a representative sex and age distribution at the first year of the time window (i.e. estimation year minus 15 years), and estimates outputted from Model 1 to probabilistically assign each individual to a specific working hours category (*i*) at the first year. Using transition probabilities outputted in Model 2, for each year over the entire time window, transitions from one working hour category to another were stochastically modelled to estimate each synthetic individual's working hours category in each year. Using Input Data 4, at each year step from the first year of the time window to the estimation year, each individual is stochastically assigned to the states of "died" or "alive". All synthetic individuals that reach the state "died" before the estimation year are censored. Using this microsimulation method, $proportion_k$ is derives using the following model (Model 3):

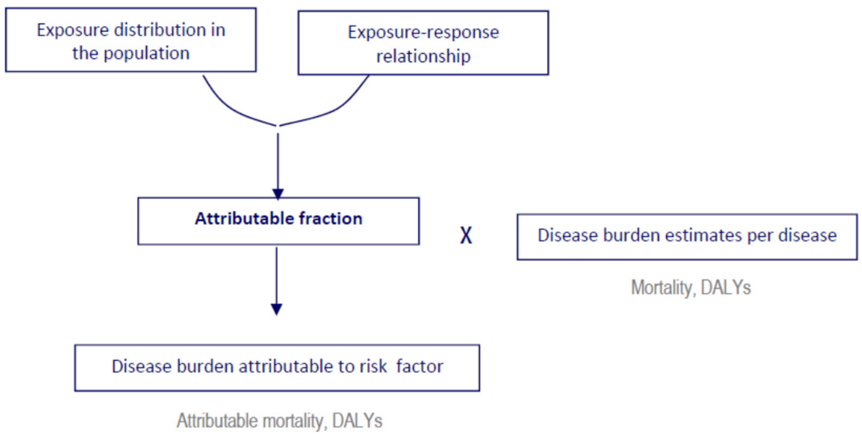
$$proportion_k = \frac{\sum_{l=1..n} \delta_{k,\max(S_l)}}{n}$$

where the summation runs through individuals "alive" in the estimation year; δ_k is the Kronecker delta function; S_l is the sequence of the l^{th} individual of all working hour categories (*i*) in each year in the time window; and where max denotes that the highest *i* that the l^{th} individual experiences in the sequence is assigned.

Model 4: Burden of disease estimation model

The Comparative Risk Assessment framework (Ezzati 2020) is used to estimate the burden of disease attributable to exposure to long working hours. We estimate the proportional reduction in death or disease that would occur if exposure was reduced to a level with a minimum risk (i.e. working 35-40 hours/week), while other conditions remain unchanged. Information on the population distribution of exposure to the risk factor is combined with information on the increased risk of acquiring or dying from the disease that was caused by exposure to the risk factor (Figure 1).

Figure 1: Comparative Risk Assessment method for burden of disease estimation



Using estimates outputted from Model 3 and Input Data 5 and 6, we calculated he "population attributable fraction" (PAF), the proportion of health outcome from the disease seen in a given population that can be attributed to exposure to the specific occupational risk factor, using Model 4:

$$PAF = \frac{\sum_{k=1}^{n} P_k(RR_k - 1)}{\sum_{k=1}^{n} P_k(RR_k - 1) + 1}$$

where P_k is the proportion of the population in a working hour category k; RR_k is the relative risk for the working hour category k; and n is the total number of long working hour categories. Applying this fraction to the total disease burden, gives the attributable disease burden.

References

Descatha et al (2018). Environ Int;119:366-378.
Descatha et al (under review). Environ Int.
Eurostat (2020). Labour market flow statistics explained. Luxembourg, Eurostat. https://ec.europa.eu/eurostat/statistics-explained/index.php/Labour_market_flow_statistics_in_the_EU
Ezzati et al (2002). Lancet;360(9343):1347-60.
Li et al (2018). Environ Int;119:558-569.
Li et al (under review). Environ Int.
UN (2019). World Population Prospects 2019. New York, UN. https://population.un.org/wpp/
WHO (2020). Life tables. Geneva, WHO. https://www.who.int/gho/mortality_burden_disease/life_tables/life_tables/en/
WHO (2018). Global Health Estimates 2016. Geneva, WHO. http://www.who.int/gho/mortality_burden_disease/en/index.htmlW