WP6: Smoking

Report on data collection for smoking and related relative risks

Esteve Fernández, MD, PhD
Jose Maria Martínez-Sánchez, BSc, MPH
Catalan Institute of Oncology

Submission date:
September 10, 2010

Table of contents

Abstract

This DYNAMO-HIA report relates to workpackage 6, “WP8: Smoking” and summarises the methods used to obtain age- and sex-specific data on smoking in EU countries with available data, as well as age- and sex-specific relative risks of the target diseases of the project by smoking status. It also provides the final estimates that will be used in the DYNAMO-HIA model.

Acknowledgements

DYNAMO-HIA is funded by the Executive Agency for Health and Consumers (EAHC, formerly known as PHEA) as part of the EU Public Health Program 2003-2008 of the European Commission's Directorate General for Health and Consumer Affairs (DG SANCO), with co-financing from the Erasmus Medical Center Rotterdam, the Institute of Public Health and the Environment in the Netherlands, the Catalan Institute of Oncology in Barcelona, the International Obesity task force, the London School for Hygiene and Tropical Medicine, the Haughton Institute in Dublin, and the Istituto Nazionale dei Tumori in Milan.

This report mainly used original data from Health Interview Surveys (HIS) from different European countries available from the Eurothine project funded by the European Commission and coordinated by Erasmus University. In this project, the investigators collected available data from several HIS conducted in18 countries and harmonised the data files. The researchers of DYNAMO-HIA WP6-Smoking asked the Principal Investigator of Eurothine to use the files produced by Eurothine and signed an agreement of use of the smoking data.For Spain, data from the 2006 National Health Interview Survey conducted by the Spanish Ministry of Health was also obtained from the public website of the Spanish Ministry of Health.

We would like specially to thank Wilma Nusselder for her guidance in several aspects of the project, including the estimation of the relative risks by time since quitting. We also thank the help of Maartje Schaap in relation to the Eurothine project data and Mónica Pérez-Ríos for the help with the search of the relative risks and interpretation of the data from the Cancer Prevention Study-II.

Introduction

DYNAMO-HIA (DYNAmic MOdel for Health Impact Assessment) project (http://www.dynamo-hia.eu/root/o14.html) is an EU funded project aiming to develop a web-based tool to assess the health impact of policies in the European Union (EU) through their influence on health determinants, including alcohol consumption. This document provides information on the project's 6th workpackage: "WP6: Smoking". It describes the sources of data that were used to deliver the required age- and sexr-specific data on tobacco smoking as well as estimates of age- and sex-specific relative risks for the selected target diseases which have been related to smoking.

WP6 was led by the Catalan Institute of Oncology but it also involved all associated partners and all 25 collaborating partners. The three main objectives of WP6 were:

1. To contribute to the discussion on specification of the model and specification of scenarios in WP4 ("Model specification and scenarios);
2. To deliver: (a) age- and sex-specific data on consumption of smoking in as many EU countries as possible, using existing publicly available data sources; and (b) age- and sex-specific relative risks (RRs) of disease incidence associated with smoking;

   This information provides input for the DYNAMO-HIA model and so links WP6 to the diseases investigated in WP9 (Cardiovascular disease and diabetes) and WP10 (Cancer);
3. To write a paper on an application of the model (this will contribute to WP2 – "Dissemination of the results").

The two main outputs of this WP are a set of data on smoking (prevalence of smokers, former smokers and never smokers, prevalence of smokers by single age, and proportion of former smokers by years since quitting) and its associated relative risks (RR).

Part 1 Estimating data on smoking

1. Choice of exposure variable

Smoking behaviour is assessed at the population level by specific tobacco surveys or, in most countries, by general Health Interview Surveys. National prevalence estimates are mostly base on self-reports of personal behaviours. Although self-reports may be subject to some misclassification bias, it is widely admitted and used that data from surveys offer a good estimate of smoking at the population level. There are several ways to ask the participants about smoking behaviour in a survey, and the operational definition of "smoker", "former smokers" and "never smokers" may vary depending on the questions included in the questionnaire. The definition of "smoker" can be based in the total number of cigarettes smoked in the respondent lifetime (usually, "have you smoked 100 cigarettes or more in your lifetime"?") or in a question about current use (usually, "do you currently smoke cigarettes"). Both definitions are however equivalent and have been combined in several systematic reviews of the literature. Time since quitting is a variable often used in the evaluation of risk linked to smoking and is collected in several ways in the surveys. Generally, age at starting smoking is asked for both current and former smokers, and in former smokers specifically, age at cessation, too. This information together with the date of the survey and the current age of the interviwee allows to derive the time (generally in years) since quitting.

In this WP, we have used the definition of "smoker" according to the original questions made in the different Health Interview Surveys, as will be reported later. The operational definitions of smoking behaviour used follow:

Table 1. Smoking behaviour operational definitions

Smoker	A person who occasionally or currently smokes at least one cigarette per day or has smoked at least 100 cigarettes during his/her life and currently smokes
Former smoker	A person who currently does not smoke but has has smoked at least one cigarette per day or has smoked 100 cigarettes during his/her life
Never smoker	A person who has never smoked or who has smoked less than 100 cigarettes during his/her life
Years since quitting	Years elapsed between quitting smoking and year of interview

2. General approach for obtaining data on smoking and evaluation of usefulness

The main source of data would be the corresponding National Health Interview Survey (NHIS) for each country, provided that these NHIS:

- Use a similar and standardized methodolgy of interview (face-to-face or telephone)
- Collect basic sociodemographic data including date of born or age at time of interview and sex
- Collect sufficient data on smoking:
    - Current smoking behaviour classified as never/former/current smoking (differences in the definition of current smoking can exist: some surveys define a smokers as the person who currently smokes at least one cigarette per day, other surveys use the definition based on having smoke at least 100 cigarettes in life and currently being smoking)
    - Data on age at starting (for both current and former smokers)
    - Data on age at stopping (for former smokers)

After reviewing the literature and based on the researchers and the DYNAMO time experience we identified the following data sources:

2.1. World Health Interview Survey

The World Health Survey (WHS) is an initiative of WHO and by now has been completed in 70 countries and the data sets have been cleaned and weighted and prepared for analysis. The field work of the WHS was conducted between 2002 to date, depending on each country.

The microdata along with all the metadata is available upon request from WHO. For the WHO European Region (EURO) data from the following countries is available: Austria, Belgium, Bosnia and Herzegovina, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Georgia, Germany, Greece, Hungary , Ireland, Israel, Italy,

Kazakhstan, Latvia, Luxembourg, Netherlands, Norway, Portugal, Russian Federation, Slovakia, Slovenia, Spain, Sweden, Turkey, Ukraine, and the United Kingdom.

The WHS questionnaire includes a section on tobacco consumption[1]. However, the data recalled is scarce, and permits only to identify smokers and non-smokers, without difference between never smokers and former smokers. Figure 1 shows the question as present in the WHS questionnaire.

Figure 1. Question on tobacco consumption in the World Health Survey.

4000. RISK FACTORS **Time Begin: __ __ : __ __**

Tobacco

(Show Tobacco list to respondent ----see Appendix A4.1)

Q4000	Do you currently smoke any tobacco products such as cigarettes, cigars, or pipes?	**1.** Daily	**2.** Yes, but not daily	**5.** No, not at all	If 2 or No: Go to Q4010
Q4001	For how many years are you smoking daily?				

On average, how many of the following products do you smoke each day?

Q4002	Manufactured cigarettes	
Q4003	Hand-rolled cigarette	
Q4004	Pipefuls of tobacco	
Q4005	Other:	

Source: World Health Survey questinnaire. World Health Organization. Available at http://www.who.int/healthinfo/survey/instruments/en/index.html

Hence, although the WHS assures a high level of availability of data from different countries and of comparability between them for many health items, its use for the purposes of the WP6 (smoking) for the DYNAMO project is not possible given the lack of identification of former smokers aside non-smokers.

[1] World Health Survey questinnaire. World Health Organization. Available at http://www.who.int/healthinfo/survey/instruments/en/index.html

2.2. Eurothine project databases

Some NHIS are available from the Eurothine project, with the necessary information as harmonized and standardized by the Eurothine investigators, a EU funded project coordinated by Erasmus MC.

The Eurothine project collected microdata from existing National Health Interview Surveys in 19 European countries as shown in Table 2.

Although the methodology of the surveys may differ from country to country, all the surveys allow to distinguish between current, former and never smokers. Moreover, the databases contain the variables indicating age at starting and age at quitting smoking.

As already checked by the Eurothine investigators the level of comparability of the smoking status and age at starting and quitting smoking is very high. The data can be used to compare countries and hence can be used to derive the indicators need for the DYNAMO project.

Table 2. Overview of national surveys included in the Eurothine project.

Country	Abbr.	Code	Name of surveys	Years	N
Norway	NOR	1	Norwegian Survey of Living Conditions	02	6820
Sweden	SWE	2	Swedish Survey of Living Conditions	00/'01	11484
Finland	FIN	3	Finbalt Health Monitor	94/'98/'00/'02/'04	20371
Denmark	DEN	4	Danish Health and Morbidity Survey 2000	00	16690
Estonia	EST	5	Health Behavior among Estonian Adult Population	02/'04	4376
Lithuania	LIT	6	Finbalt Health Monitor	94/'98/'00/'02/'04	11647
Latvia	LAT	7	Finbalt Health Monitor	98/'00/'02/'04	8488
Czech Rep	CZR	8	Sample Survey of the Health Status of the Czech Population	02	2476
Slovakia	SLO	9	Health Monitor Survey 2002	'02	1528
England	ENG	10	English Health Survey 2001	'01	15767
Ireland	IRE	11	Living in Ireland Panel Survey	95/'02	15051
Netherlands	NET	12	Permanent Onderzoek Leefsituatie (POLS)	03/'04	15803
Belgium	BEL	13	Health Interview Survey 1997 + 2001	97/'01	18481
Germany	GER	14	German National Health Examination and Interview Survey	'98	7124
France	FRA	15	2004 Health, Health Care and Insurance Survey (IRDES)	'04	17828
Italy 1	ITA_1	16	Health and health care utilization 1999-2000	99/'00	118245
Italy 2	ITA_2	16	Multipurpose Family Survey. Aspects of daily living.	'00	49373
Spain	SPA	17	National Health Survey 2001	'01	20748
Portugal	POR	18	National Health Survey (1998/1999)	98/'99	40917
Hungary	HUN	19	National Health Interview Survey Hungary	00/03	10532
Total					388435

Source: Schaap et al. Specification of data files created within the EUROTHINE project. Harmonized files based on National Health Interview Surveys. Rotterdam: Erasmus MC; 2006.

2.3. Eurobarometer on smoking

The current Flash Eurobarometer on *Tobacco* (Flash No 253), requested by the Directorate General Health and Consumers, was conducted to evaluate the EU citizens' and Norwegians' attitudes towards tobacco. The analytical report with the main results by country and for the EU has recently been published[2].

The survey examined smoking habits and consumption of non-combustible tobacco products, exposure to tobacco smoke at home and at the workplace, and other related topics on attitudes towards smoking restrictions, health warnings, and purchasing of tobacco products.

The Flash Eurobarometer on Tobacco (Flash No 253) fieldwork was conducted between 13 and 17 December 2008. Over 26,500 randomly-selected citizens aged 15 years and over were interviewed in the 27 EU Member States and in Norway. Interviews were predominantly carried out via fixed-line telephone, reaching approximately 1000 EU citizens in each country (in Cyprus, Luxembourg and Malta the targeted size was 500). Parts of the interviews in Finland, Austria, Portugal and Italy were conducted over mobile telephones. Due to the relatively low fixed-line telephone coverage in Bulgaria, the Czech Republic, Estonia, Latvia, Lithuania, Hungary, Poland, Romania and Slovakia, 300 individuals were sampled and interviewed on a face-to-face basis.

The Flash Eurobarometer on Tobacco (Flash No 253) allows to distinguish between current (daily and occasional), former, and never smokers. However, the questionnaire did not include specific questions on age at starting nor age at quitting (Figure 2).

[2] European Commision. The Gallup Organisation. Survey on Tobacco. Analytical report. Flash Eurobarometer 253. Brussels: European Commision; 2009. Available at: http://ec.europa.eu/health/ph_determinants/life_style/Tobacco/keydo_tobacco_en.htm

Figure 2. Question on tobacco consumption in the Flash Eurobarometer on Tobacco 2008 (Flash No 253).

III. Questionnaire

Smoking behaviour

Q1. Regarding smoking cigarettes, cigars or a pipe, which of the following applies to you?

[READ OUT – ONE ANSWER ONLY]	
You smoke every day	1
You smoke occasionally	2
You used to smoke but you have stopped	3
You have never smoked	4
[DK/NA]	9

Source: European Commision. The Gallup Organisation. Survey on Tobacco. Analytical report. Flash Eurobarometer 253.

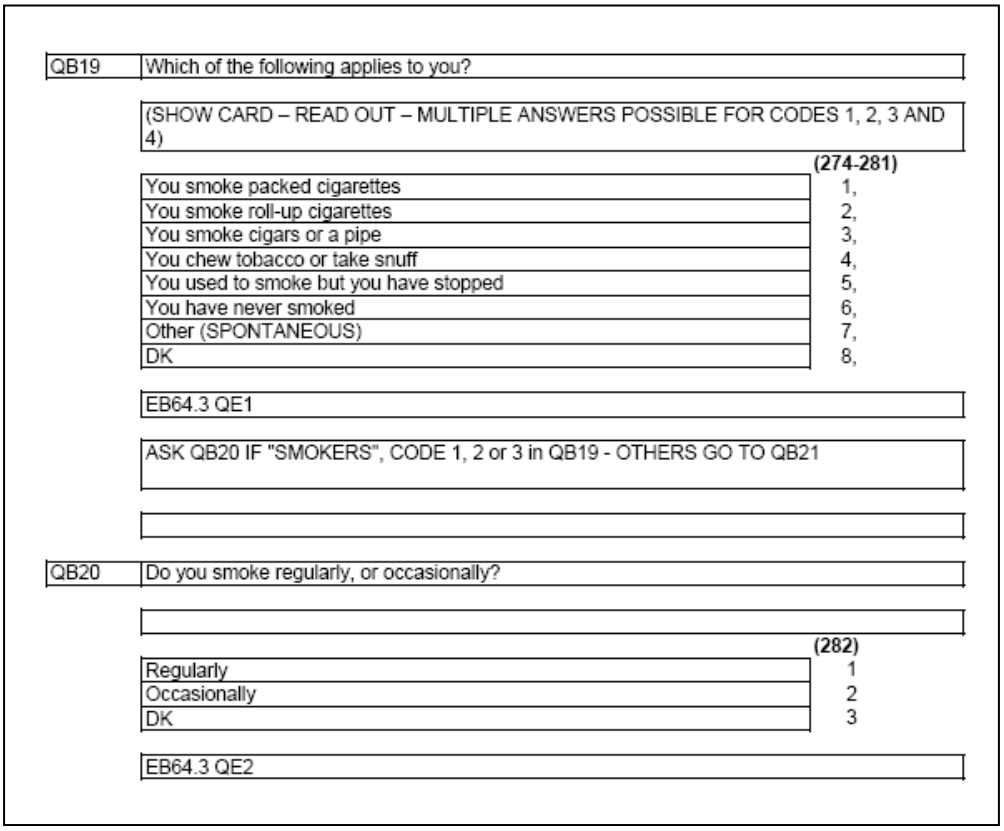
Although the Eurobarometer assures a high level of availability of data from all EU countries plus Norway and of comparability between them, its use for the purposes of the WP6 (smoking) for the DYNAMO project is not possible given the lack of information that permits computation of the prevalence of former smokers by time since quitting. Moreover, the small sample size in each country (approx. 1000 respondents in each country except Cyprus, Luxembourg and Malta with 500 respondents) would have make difficult to obtain accurate and reliable estimates of former smoking according to time since quitting. With prevalence of former smokers around 20% (as in many countries), there would be approximately 200 former smokers to split according to time since quitting.

Previous Eurobarometers on smoking (years 2005 and 2006)[3,4] did unfortunately in a similar way, with the questionnaire allowing to distinguish between current, former and never smokers (Figures 3 & 4) but without information on age or years since quitting.

[3] European Commision. TNS Opinion & Social. Attitudes of Europeans towards tobacco. Special Eurobarometer 239. Brussels: European Commision; 2006. Available at: http://ec.europa.eu/health/ph_determinants/life_style/Tobacco/keydo_tobacco_en.htm

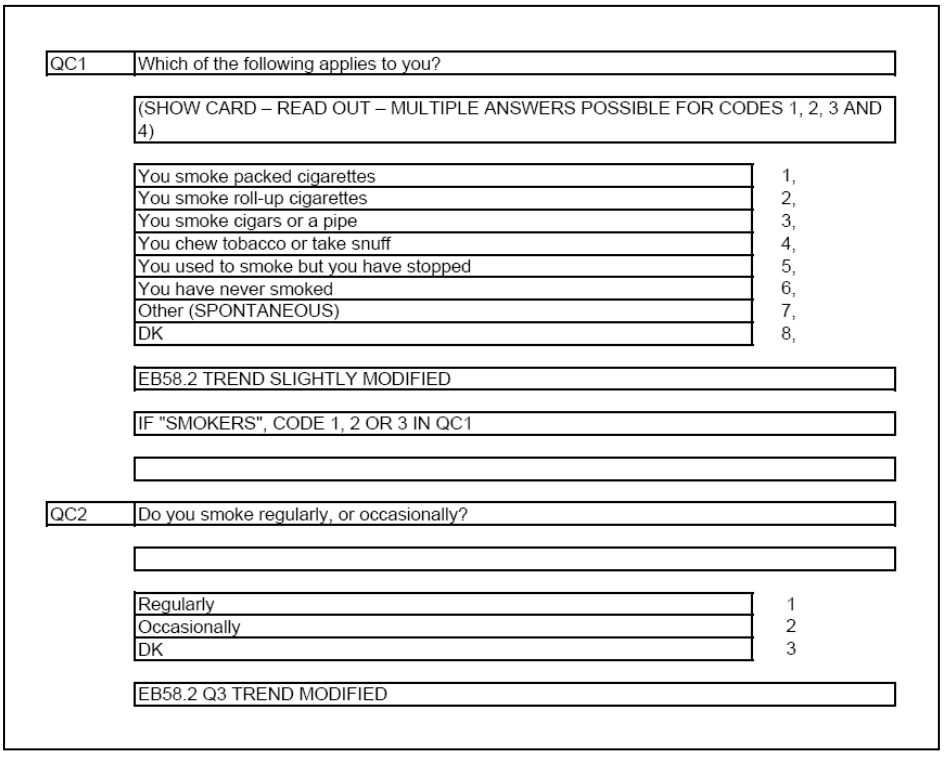
[4] European Commision. TNS Opinion & Social. Attitudes of Europeans towards tobacco. Special Eurobarometer 272c. Brussels: European Commision; 2007. Available at: http://ec.europa.eu/health/ph_determinants/life_style/Tobacco/keydo_tobacco_en.htm

Figure 3. Question on tobacco consumption in the Special Eurobarometer 239 (2005).



QB19	Which of the following applies to you?	
	(SHOW CARD – READ OUT – MULTIPLE ANSWERS POSSIBLE FOR CODES 1, 2, 3 AND 4)	
		(274-281)
	You smoke packed cigarettes	1,
	You smoke roll-up cigarettes	2,
	You smoke cigars or a pipe	3,
	You chew tobacco or take snuff	4,
	You used to smoke but you have stopped	5,
	You have never smoked	6,
	Other (SPONTANEOUS)	7,
	DK	8,
	EB64.3 QE1	
	ASK QB20 IF "SMOKERS", CODE 1, 2 or 3 in QB19 - OTHERS GO TO QB21	
QB20	Do you smoke regularly, or occasionally?	
		(282)
	Regularly	1
	Occasionally	2
	DK	3
	EB64.3 QE2	

Figure 4. Question on tobacco consumption in the Special Eurobarometer 272c (2006).



QC1	Which of the following applies to you?	
	(SHOW CARD – READ OUT – MULTIPLE ANSWERS POSSIBLE FOR CODES 1, 2, 3 AND 4)	
	You smoke packed cigarettes	1,
	You smoke roll-up cigarettes	2,
	You smoke cigars or a pipe	3,
	You chew tobacco or take snuff	4,
	You used to smoke but you have stopped	5,
	You have never smoked	6,
	Other (SPONTANEOUS)	7,
	DK	8,
	EB58.2 TREND SLIGHTLY MODIFIED	
	IF "SMOKERS", CODE 1, 2 OR 3 IN QC1	
QC2	Do you smoke regularly, or occasionally?	
	Regularly	1
	Occasionally	2
	DK	3
	EB58.2 Q3 TREND MODIFIED	

Source: European Commision. TNS Opinion & Social. Attitudes of Europeans towards tobacco

2.4. Other sources

The access to other NHIS was also contemplated in order to obtain more recent data or data from countries not included in the Eurothine:

- Spain: The Spanish NHIS from 2006 was obtained from the Ministry of Health to update the 2001 data available from the Eurothine project

- Poland: requests of the availability of data were made to different contacts in Poland with no success
- Italy: request of microdata from a more ecent NHIS was done but the data was not available. Alternative source of data (annual DOXA survey on tobacco consumption) for the year 2006 was available but disregarded since the sample size was smaller than the 2000 NHIS data already available from Eurothine
- Portugal: requests of the availability of data were made to different contacts in Portugal but no update of the NHIS was available

3. Data collection and estimation methods

3.1Criteria used for selecting sources of individual level data

The microdata available from the Eurothine project met the following criteria:

Time frame

- The study was relatively recent: surveys conducted on or since 1997, most studies were conducted around 2000.

Study design and sample

- Population-based cross-sectional studies
- The reference population was described and corresponded as closely as possible to the national population (thus regional surveys and those of special interest or of particular groups were excluded)

- The sampling strategy was as close as possible to random sampling
- The sample was representative of the reference population
- The sample size was large
- As wide an age range as possible (from 15 years onwards) was included
- Data were available by age and sex
- The level of non-response was documented

Validity of the methods

- The methods used to collect data were as free of bias as possible.
- Data were collected at the level of the individual.

Type of information

- Data on smoking had to be available according to the set categories defined under Section 1 above, stratifying by sex and age group.

3.2 Characteristics of included and excluded individual-level data

Individual-level estimates of smoking were obtained for 18 EU countries. Details of variables available for each country are described in Table 3. All were from national surveys and previously collected by the Eurothine project except updated data from Spain. All the studies covered the period 2000 to 2006 with the exception of the NHIS from Germany and Portugal (table 1). When data sets from different editions of the survey were available (eg, Lithuania) the most recent was used. When two surveys in consecutive years (eg, years 2000 and 2001, or years 2002 and 2004) were available and the sample size was relatively small, we merged the datasets to obtain a greater sample size (eg, Sweden, years 2000+2001). We therefore included the datasets from those countries with information on the prevalence of current smokers and enough information to compute the years since quitting smoking (table 4).

Thus, from the 18 countries with datasets available, only 8 countries have enough data to compute all the target variables for the DYNAMO including time since quitting

(table 3). Basic information on the prevalence of smokers, former smokers, and never smokers was however available for 28 EU countries from the Eurobarometer Survey (aggregated data as published by the European Commission).

Table 3. Availability of data from 28 EU countries

	% smokers	% ex smokers	% never smokers	% of smokers by single year of age	% of ex smokers by years since quitting
Belgium	x	x	x	x	
Bulgaria	x	x	x		
Czech Rep.	x	x	x	x	
Denmark	x	x	x	x	
Germany	x	x	x	x	x
Estonia	x	x	x	x	
Greece	x	x	x		
Spain	x	x	x	x	x
France	x	x	x	x	x
Ireland	x	x	x	x	
Italy	x	x	x	x	x
Cyprus	x	x	x		
Latvia	x	x	x	x	
Lithuania	x	x	x	x	
Luxembourg	x	x	x		
Hungary	x	x	x	x	x
Malta	x	x	x		
Netherlands	x	x	x	x	x
Austria	x	x	x		
Poland	x	x	x		
Portugal	x	x	x	x	x
Romania	x	x	x		
Slovenia	x	x	x		
Slovakia	x	x	x	x	
Finland	x	x	x	x	
Sweden	x	x	x	x	
United Kingd	x	x	x	x	x
Norway	x	x	x	x	

Table 4. Available datasets with micro-data from the Eurothine project

Country	Year of survey(s)
Norway	2002
Sweden	2000+2001
Finland	2002+2004
Denmark	2000
Estonia	2002+2004
Lithuania	2002+2004
Latvia	2002+2004
Czech Republic	2002
Slovakia	2002
England	2001
Ireland	2002
Netherlands	2003+2004
Belgium	2001
Germany	1998
Italy	2000
France	2004
Portugal	1998+1999
Hungary	2000+2003

3.3 Estimation of individual-consumption from survey data

Data from individual-based surveys were re-analysed by us using the harmonised datasets provided by the Eurothine time. In each data set the following variables were selected:

- ID of country
- Sex
- Age
- Data on smoking:
  - Current smoking behaviour classified as never/former/current smoking
  - Age at stopping (for former smokers)
- Country code [variable name: COUNTRY]
- File part [variable name: PART]
- Year of the survey [variable name:YEAR]
- Selection weight (for DK, IRE, GER, NET) [variable name: W_SEL]
- Sex [variable name: SEX]
- Age in continous form [variable name: AGE]
- Age category [variable name: AGECAT5]
- Smoking status [variable name: SMOKSTAT]
- Age at smoking cessation [variable name: SMOKQUIT]

3.4. Imputation of prevalence data by single year of age

While the datasets provided us with estimates by sex and age group, the DYNAMO-HIA model requires data by one-year age interval up until 95 years of age. It was thus decided to smooth the estimates of smoking with age in order to avoid important and unrealistic gaps at the junction of the age groups. Also, because several surveys considered a limited age range (e.g. no data below 20 years of age or above 65 years), it was assumed that smoothed data would provided better estimates of smoking for those above or below the survey age range.

In order to smooth the prevalence estimates, we proceeded as follows:

1) We computed the prevalence of smoking (or of former smoking) by 10-year age-groups (e.g. 20-29 years) except the first age-group (16-19 years) and further imputed the single year of age prevalence following a standard linear interpolation method

2) To compute the last or lasts single year of age prevalence rates (ie, from 90 to 99 years, or from 89 to 99 years), we assumed a prevalence for 99 years old class attending the prevalence of the last group available, the previous trend observed and the prevalence of a country with similar sociodemographic characteristics and development of the tobacco epidemic. Hence, prevalence by single year of age were interpolated from the last available figure computed from the database to the assumed figure.

Part 2 Estimating risk factor-disease relationships

1. Choice of outcomes

Within this project, a limited number of health outcomes were selected to be modelled separately using the DYNAMO-HIA model. These health outcomes were selected based on two criteria: 1) best evidence of a risk factor-disease relationship for most risk factors examined in the project, i.e. alcohol consumption, smoking, and overweight/obesity; and 2) the prevalence of the disease. Effects of each risk factor through diseases not modelled separately can be included in the DYNAMO-HIA model by using relative risks (RRs) for all-cause mortality and all-cause disability.

The selected health outcomes discussed here include:

- all-cause mortality;
- ischaemic heart disease (IHD);
- stroke;
- diabetes mellitus;
- chronic obstructive pulmonary disease (COPD);
- cancer of the lung;
- cancer of the colon and rectum;
- cancer of the mouth and oropharynx; and
- breast cancer

For smoking, the choice of outcomes was guided mainly by evidence of causality given by two reports from the US Surgeon General and one IARC Monograph:

- The Health Benefits of Smoking Cessation. A Report of the Surgeon General. Year 1990
- The Health Consequences of Smoking. A Report of the Surgeon General. Year 2004
- Tobacco Smoke and Involuntary Smoking. IARC Monograph on the Evaluation of Carcinogenic Risks to Humans. Vol. 83. Year 2004.

2. Overview of the association of target DYNAMO diseases to smoking behaviour

Both the Surgeon General Reports and the IARC Monograph consider that there is sufficient evidence in humans to infer a causal relationship between smoking and:

- Lung cancer
- Oral cavity and pharyngeal cancer
- Esophageal cancer

These reports, however, conclude that the evidence is suggestive of no causal relationship between active smoking and breast cancer.

The IARC Monograph states that "it is not possible to conclude that the association between tobacco smoke and colorectal cancer is causal". The 2004 Surgeon General states that "the evidence is suggestive but not sufficient to infer a causal relationship between smoking and colorectal adenomatous polyps and colorectal cancer".

Regarding endometrial cancer, the IARC Monograph says that "there is evidence suggesting lack of carcinogenicity of tobacco smoking in humans for cancer of the female breast and the endometrium" whereas the 2004 Surgeon General Report considers that "the evidence is sufficient to infer that current smoking reduces the risk of endometrial cancer in postmenopausal women"

The non-neoplastic diseases are considered by the 2004 Surgeon General Report:

- "The evidence is sufficient to infer a causal relationship between smoking and coronary heart disease."
- "The evidence is sufficient to infer a causal relationship between smoking and stroke."
- "The evidence is suggestive to infer a causal relationship between active smoking and chronic obstructive pulmonary disease morbidity and mortality".

Diabetes has not been included in the 2004 Surgeon General Report (nor in the 1990 Report) given the lack of previous evidence linking diabetes to tobacco smoking. In addition, we have performed a specific bibliographic search in the PubMed and Web of Knowledge databases, in search of systematic reviews and/or meta-analysis assessing the causal relationship between tobacco smoking and diabetes. We find a relevant meta-

amalysis that included 25 prospective cohort studies with a pooled RR of 1.44 (95% CI: 1.31-1.58). Several factors, however, prevent from considering smoking as a true risk factor for diabetes. The studies included in the meta-analysis are heterogenous and as socioeconomic status, education, obesity and physical activity cannot be disregarded as potential confounders. Finally, this systematic review might be affected by publication bias. In view of this criticisms, it is clear that the meta-analysis provides evidence of an association between smoking and diabetes, but this associatin is not causal. Therefore, we decided not to include diabetes as a target disease for smoking.

3. Data collection and estimation methods

As derived from the previous point, the relative risks (RR) of colorectal cancer, breast cancer, endometrial cancer, and diabetes according to smoking status are not of interest for the model.

We have reviewed the Surgeon General Reports and IARC Monograph to obtain summary RRs from meta-analysis of the existing studies. However, the heterogeneity of the studies prevented the meta-analysis in these reports. It provides extensive summary tables with the main results of the studies. These tables mostly include the RR (or in some cases the absolute rates) of mortality, and only the RR of the incidence of the target DYNAMO diseases for a few studies. Most studies have had as primary outcome mortality instead of the inicdence of the disease. However, in most of the diseases considered, the incidence rate to mortality rate ratio is greater than 0,7, and in some cases such as lung cancer, oesophageal cancer, and COPD approaches to 1, thus indicating that the mortality is a good estimate of incidence.

These RRs are provided for current versus never smokers, and in some cases also the RR of mortality of former smokers (and in some studies by amount of cigarettes smoked, and/or time since quitting or age at quitting). The detail of the tables is poorer in the 2004 Surgeon Report than in the previous 1990 Surgeon General Report or the 2004 IARC Mongraph. It is also important to note the variability in the categorization of time since quitting across studies.

Regarding the three cancer sites (lung, mouth/pharynx, oesophageal) linked to tobacco smoking, the IARC Monograph included 45 cohort studies and a higher number of case-control studies (not considered in this overview):

- Lung cancer: 21 cohort studies with RR or SMR of lung cancer (death) for former smokers. 11 of these have data on time since quitting.
- Mouth and pharynx cancer:11 cohort studies with 4 studies with data on time since quitting
- Oesophageal cancer: 44 cohort studies, 26 of them with data on former smoking.

After carefully reviewing these Reports and Monographs, and considering the quality of the studies and also its previous use, we were inclined to use the RR of death derived from the second Cancer Prevention Study (CPS-II) conducted by the American Cancer Society among a cohort of one million American volunteers followed-up during 6 years (1982 to 1988). The RR derived from CPS are *de facto* the "standard" for computation of attributable mortality to smoking not only in the USA but in other countries, and have also been used in established models such as the SAMMEC (Smoking-Attributable Mortality, Morbidity, and Economic Costs) project of the US CDC.

These RR derives from the CPS-II study as provided by the American Cancer Society but have not been published. They are available from different sources, for example they are included in the 2004 Surgeon General Report and also in the SAMMEC webpage.

Tables 5 shows the RR by smoking status (never, current, former) taking as reference group the "never smokers". CPS-II does not provide the RR of death for IHD, stroke, and "all causes" for all ages combined.

Table 6 shows the available RR for former smokers by age groups. Since no RR were available for some diseases (mouth and pharynx cancer, oesophageal cancer, IHD, cerebrovascular disease, and "all causes"combined) we have assigned the available RR for each age-group.

Table 5. RR of death (DYNAMO target diseases) according to smoking status by sex

	Men			Women		
Disease Category	**Never Smoker**	**Current Smoker**	**Former Smoker**	**Never Smoker**	**Current Smoker**	**Former Smoker**
Lip, Oral Cavity, Pharynx	1 (reference)	10.89	3.40	1 (reference)	5.08	2.29
Oesophagus	1 (reference)	6.76	4.46	1 (reference)	7.75	2.79
Endometrium		*not sufficient evidence to conclude a causal association*				
Trachea, Lung, Bronchus	1 (reference)	23.26	8.70	1 (reference)	12.69	4.53
Colon		*not sufficient evidence to conclude a causal association*				
Breast		*not sufficient evidence to conclude a causal association*				
Ischemic Heart Disease						
Persons Aged 35–64	1 (reference)	2.80	1.64	1 (reference)	3.08	1.32
Persons Aged 65+	1 (reference)	1.51	1.21	1 (reference)	1.60	1.20
Cerebrovascular Disease						
Persons Aged 35–64	1 (reference)	3.27	1.04	1 (reference)	4.00	1.30
Persons Aged 65+	1 (reference)	1.63	1.04	1 (reference)	1.49	1.03
Chronic Airway Obstruction	1 (reference)	10.58	6.80		13.08	6.78
Diabetes		*not sufficient evidence to conclude a causal association*				
All causes	1 (reference)	2.07	1.35	1 (reference)	1.74	1.23

Source CPS–II(82-88): Unpublished estimates provided by American Cancer Society (ACS). See Thun MJ, Day-Lally C, Myers DG, et al. Trends in tobacco smoking and mortality from cigarette use in Cancer Prevention Studies I (1959 through 1965) and II (1982 through 1988). In: Changes in cigarette-related disease risks and their implication for prevention and control. Smoking and Tobacco Control Monograph 8. Bethesda, MD: US Department of Health and Human Services, Public Health Service, National Institutes of Health, National Cancer Institute 1997;305–382. NIH Publication no. 97–1213. https://apps.nccd.cdc.gov/sammec/show_risk_data.asp

Table 6. RR of death (DYNAMO target diseases) for former smokers by age and sex

	Male			Female		
Disease Category	**Never Smoker**	**Current Smoker**	**Former Smoker**	**Never Smoker**	**Current Smoker**	**Former Smoker**
Lip, Oral Cavity, Pharynx						
Persons Aged 35–39	1 (reference)	10.89	3.40	1 (reference)	5.08	2.29
Persons Aged 40-44	1 (reference)	10.89	3.40	1 (reference)	5.08	2.29
Persons Aged 45–49	1 (reference)	10.89	3.40	1 (reference)	5.08	2.29
Persons Aged 50–54	1 (reference)	10.89	3.40	1 (reference)	5.08	2.29
Persons Aged 55–59	1 (reference)	10.89	3.40	1 (reference)	5.08	2.29
Persons Aged 60-64	1 (reference)	10.89	3.40	1 (reference)	5.08	2.29
Persons Aged 65+	1 (reference)	10.89	3.40	1 (reference)	5.08	2.29
Oesophagus						
Persons Aged 35–39	1 (reference)	6.76	4.46	1 (reference)	7.75	2.79
Persons Aged 40-44	1 (reference)	6.76	4.46	1 (reference)	7.75	2.79
Persons Aged 45–49	1 (reference)	6.76	4.46	1 (reference)	7.75	2.79
Persons Aged 50–54	1 (reference)	6.76	4.46	1 (reference)	7.75	2.79
Persons Aged 55–59	1 (reference)	6.76	4.46	1 (reference)	7.75	2.79
Persons Aged 60-64	1 (reference)	6.76	4.46	1 (reference)	7.75	2.79
Persons Aged 65+	1 (reference)	6.76	4.46	1 (reference)	7.75	2.79
Lung cancer						
Persons Aged 35–39	1 (reference)	1.30	1	1 (reference)	2	1
Persons Aged 40-44	1 (reference)	1	1	1 (reference)	1	1
Persons Aged 45–49	1 (reference)	5.78	2.37	1 (reference)	18.08	8.07
Persons Aged 50–54	1 (reference)	24.97	10.70	1 (reference)	11.14	3.28
Persons Aged 55–59	1 (reference)	34.02	11.66	1 (reference)	17.87	5.33
Persons Aged 60-64	1 (reference)	31.47	11.71	1 (reference)	13.32	4.91
Persons Aged 65+	1 (reference)	28.40	9.70	1 (reference)	17.49	5.54

Continues…

Ischemic Heart Disease							
Persons Aged 35–39	1 (reference)	3.25	1.21		1 (reference)	1	1.44
Persons Aged 40-44	1 (reference)	4.71	1.15		1 (reference)	1.89	2.25
Persons Aged 45–49	1 (reference)	5.85	2.03		1 (reference)	7.71	2.08
Persons Aged 50–54	1 (reference)	3.69	1.93		1 (reference)	5.69	2.95
Persons Aged 55–59	1 (reference)	2.71	1.64		1 (reference)	3.06	1.19
Persons Aged 60-64	1 (reference)	2.39	1.58		1 (reference)	2.56	1.08
Persons Aged 65+	1 (reference)	1.91	1.40		1 (reference)	2.48	1.22
Cerebrovascular disease							
Persons Aged 35–39	1 (reference)	1	1		1 (reference)	2	1
Persons Aged 40-44	1 (reference)	1.05	1		1 (reference)	5.67	2.25
Persons Aged 45–49	1 (reference)	3.75	1		1 (reference)	8.22	1.19
Persons Aged 50–54	1 (reference)	6.08	2.24		1 (reference)	4.58	1.38
Persons Aged 55–59	1 (reference)	3.96	1.14		1 (reference)	5.77	1.22
Persons Aged 60-64	1 (reference)	2.55	1.01		1 (reference)	2.76	1.28
Persons Aged 65+	1 (reference)	2.69	1.29		1 (reference)	2.58	1.14
Chronic Airway Obstruction							
Persons Aged 35–39	1 (reference)	1	1		1 (reference)	1	1
Persons Aged 40-44	1 (reference)	1	1		1 (reference)	1	1
Persons Aged 45–49	1 (reference)	1	1		1 (reference)	1	1
Persons Aged 50–54	1 (reference)	8.13	3.06		1 (reference)	12.92	7.39
Persons Aged 55–59	1 (reference)	9.80	8.25		1 (reference)	9.47	5.55
Persons Aged 60-64	1 (reference)	13.21	12.65		1 (reference)	11.19	6.63
Persons Aged 65+	1 (reference)	18.93	11.92		1 (reference)	14.72	9.73
All causes							
Persons Aged 35–39	1 (reference)	2.07	1.35		1 (reference)	1.74	1.23
Persons Aged 40-44	1 (reference)	2.07	1.35		1 (reference)	1.74	1.23
Persons Aged 45–49	1 (reference)	2.07	1.35		1 (reference)	1.74	1.23
Persons Aged 50–54	1 (reference)	2.07	1.35		1 (reference)	1.74	1.23
Persons Aged 55–59	1 (reference)	2.07	1.35		1 (reference)	1.74	1.23
Persons Aged 60-64	1 (reference)	2.07	1.35		1 (reference)	1.74	1.23
Persons Aged 65+	1 (reference)	2.07	1.35		1 (reference)	1.74	1.23
	American Cance	American Cancer Society´s Cancer Prevention Study II age-specific relative risks (1982-1988).					

4. Computation of RR of the target diseases by years since quitting

The DYNAMO model needs RRs of the target diseases (in the case ofsmoking: oral cavity cancer, oesophagus cancer, lung cancer, ischemic heart disease, cerebrovascular disease, chronic airway obstruction, and all causes of mortality) by time since quitting stopping, age and sex. Since the data available from the CPS-II study is incomplete (no data for single year of age and for single year of time since quitting), an estimation of the RRs by age and time since quitting was used. We implemented the approach used by Hoogenveen et al.e As described in Appendix 3 of the paper by Hoogenveen et al., the method used to calculate relative risks of former smokers that depend on time since smoking cessation were estimated. We used the regression parameter from the distribution of all former smokers over time since cessation and mean relative risks of all former smokers:

»The statistical model is defined for the relative risks of former smokers compared to never smokers as a function of the time since smoking cessation. These relative risks comprise both all cause mortality and incidence of chronic diseases. The relative risks of former smokers decrease over time since cessation, meaning that the effect of past smoking behavior gradually disappears. We made the following assumptions:

- The relative risk of quitters equals the relative risk of current smokers.
- The relative risk of former smokers approaches the relative risk of never smokers, i.e. value 1.
- Relative risks of former smokers show a time-constant proportional decrease.
- The proportionality coefficients that describe the rate of decrease over time of the relative risks decrease proportionally over age

These assumptions result in the following formulas for the relative risk:

$$RR_{former}(a, s) = 1 + (RR_{current}(a) - 1) \exp(-\gamma(a)s)$$
$$\gamma(a) = \gamma_0 \exp(-\eta\, a^*(a))$$

[e] Hoogenveen RT, van Baal PHM, Boshuizen HC, Feenstra TL. Dynamic effects of smoking cessation on disease incidence, mortality and quality of life: The role of time since cessation. *Cost Effectiveness and Resource Allocation*. 2008; 6:1

with:

a age

a*(a) transformation of a, $a^*(a) = (a-50)^+$: the non-negative value of a-50

s time since smoking cessation

γ regression coefficient of time dependency

η regression coefficient of age dependency

$RR_{current}(a)$ relative risks of current smokers at age a

$RR_{former}(a)$ mean relative risks of all former smokers at age a

With this approach, we were able to estimate the RR of death for the 6 target diseases for smoking plus all cause mortality by single year of age and time since quitting. The results in the DYNAMO database include an 82 (age) by 20 (years since quitting) matrix with the corresponding RR for each disease.