Health Impact Assessment (HIA) – the evaluation of policies, projects, or proposals concerning their effects on human health – becomes increasingly common practice at the local, national, and EU-level. So far, no standard tool exists to aid the quantification step in HIA. This thesis proposes dynamic population health modeling as a methodological foundation for quantitative HIA by motivating and introducing a ready-to-use software tool for this purpose: DYNAMO-HIA. In addition, selected applications are presented ranging from the health consequences of an EU-wide tax increase on alcohol to the quantification of the life-long health benefits of reducing obesity when entering adulthood.

Dynamic Population Health Modeling for Quantitative Health Impact Assessment

Stefan K. Lhachimi

Dynamic Population Health Modeling for Quantitative Health Impact Assessment

Methodological Foundation and Selected Applications

Stefan K. Lhachimi

Dynamic Population Health Modeling
for
Quantitative Health Impact Assessment

Methodological Foundation and Selected Applications

Dynamisch modelleren van de volksgezondheid
voor kwantitatieve gezondheidseffectschatting
Methodologische onderbouwing en geselecteerde toepassingen

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
rector magnificus
Prof.dr. H.G. Schmidt
and in accordance with decision of the Doctorate Board
The public defence shall be held on
Friday, 18th of November 2011 at 9:30 hrs
by
Stefan Karim Lhachimi
born in Bernburg an der Saale, Germany

Doctoral Committee

Promoters:	Prof.dr. J.P. Mackenbach
	Prof.dr. H.C. Boshuizen
Other members:	Prof.dr. M. McKee
	Prof.dr. D. van de Mheen
	Prof.dr. H. van Oers
Copromoter:	Dr. W.J. Nusselder

Denn sehet die Teppichböden, sie weben nicht, sie streben nicht und dennoch reichen sie von Wand zu Wand.

W.M.

Contents

Appendix

List of Figures

List of Tables

1 Introduction

All models are wrong, some are useful.
George E. P. Box

Health impact assessment: Between technical Exercise and participatory process

The definition most widely used for health impact assessment (HIA) stems from the Gothenburg consensus paper and states that HIA is *a combination of procedures, methods, and tools by which a policy, program, or project may be judged as to its potential effects on the health of a population, and the distribution of those effects within the population.*[1]

HIA typically informs policy options outside the health sector but can also cover a dedicated health policy. The roots of HIA can be traced to two prior developments (see Veerman for a detailed overview[2]). The first root is the extension of environmental impact assessments (EIA) to include also health. Particularly in developing countries, it became increasingly apparent that ambitious infrastructure projects, such as dams or irrigation, have not only consequences for the environment but also for human health. Adding health as a dimension when conducting an EIA is now more common but not standard. The second root stems from a movement that was more concerned with the social and behavioral determinants of health such as the WHO Healthy Cities Initiative. This approach is holistic and emphasizes the interaction with affected individuals and communities. Particular attention is payed to the causes and mitigation of social inequalities in health. These two different roots and the very diverse field of application (from EU-wide policies to urban development) partly explain the variety of existing approaches

and certain tensions within the field. But all HIAs have in common that they are simply speaking "evaluations before the fact."[2]

Objectives of HIA

An HIA exercise usually has three main objectives: First, to *predict* the impact of a policy, second, to allow *participation* of stakeholders in the assessment process, and, third, to *inform* the decision making process.[3]

The first objective is close to epidemiological considerations. The causal pathway by which the policy affects health has to be identified, i.e. which risk factors are affected. When the causal chain is established, HIA aims to predict the health changes for the *whole* population and the distribution *within* the population. Not only the net effect is important, but it is also crucial to identify both winners and losers of a policy. It is necessary to be as thorough as possible and to detect the negative and the positive effects of the policy on health.

Some regard the second objective – to make the health assessment of a policy proposal a participatory endeavor – as a key feature of the HIA process. The main arguments are that residents have the right to be informed and that they are often the best source of information in an assessment. They know their community best and they are the ones who are affected by the decision the most.[4] Critics, though, argue that it is onerous to motivate community members to participate and for larger projects this becomes impracticable. Certainly, for policy decisions at the national level it will be difficult to conduct an all-inclusive process.[5] Yet, for every decision under deliberation there are different stakeholders with diverse preferences involved and every HIA must accommodate their legitimate interests by being as transparent as possible.

Third, an HIA has to inform the decision making process with applicable knowledge. The derived information has to be put into an understandable and applicable form for the specific context. Giving a plethora of data or a simple *yes/no* answer will not be convincing in the policy arena. This objective stems partly from the practical observation that the assessment is normally done by different people than the actual decision taking. Although HIA is becoming more and more common, many administrative units simply lack the adequately trained personnel to conduct an HIA.[6]

In this thesis we limit the focus on the two technical core tasks of an HIA (compare for a recent definition[2]):

- supporting decision-makers in choosing between options and
- predicting the future consequences of implementing different options.

Why numbers count

Quantification is desirable in most if not all decision situations. Even in the situation where two options are compared that are both beneficial, have the same costs, and the same chance of success, quantifying the outcome allows choosing the (even) better option. Hence, quantification allows a ranking of proposals where strictly speaking "doing nothing", i.e. postponing the decision, is always one of the options. Furthermore, if policies are discussed that yield both benefits and losses – be it to a group of individuals, be it to the population as a whole – quantification in terms of changes in health allows a more precise balancing of competing criteria. And, pragmatically speaking, recommendations without quantification have difficulties standing up in the policy arena; in particular if they are contrary to proposals that promise (quantified) economic well-being.

For some, the quantification of future health outcomes of the assessed policy proposals is not necessarily at the core of HIA as they rather stress the qualitative and participatory elements of HIA. Those are both valid elements of an HIA. Qualitative assessment of a policy proposal may reveal connections between the proposal and individual or population health which have been unknown or ignored when conceiving the policy. Similarly, the participatory elements of an HIA, i.e. consulting with affected individuals, health experts, and other stakeholders, not only helps to reveal such connections; it also allows – by having deliberative interactions – to find a better policy outcome for all involved. Nevertheless, a fair and valid quantification of all proposals under discussion will ceteris paribus only aid a rational decision process.

Structure of an HIA

HIA is a multi-stage process commonly divided into five steps[3,7–9]: *Screening* is the assessment of – optimally – any proposal or policy whether it has an intended or unintended effect on population health and, hence, should be

subject to an HIA. *Scoping* refers to the process of identifying the potential health impacts of a policy, program, or project before they are quantified. *Effect analysis* is at the core of an (extensive) HIA detailing and appraising the potential consequences for population health. *Reporting of findings* ensures that the insights gained of an analysis are being fed back into the policy process, preferably in a clear and transparent manner, informing all stakeholders sufficiently. *Monitoring and evaluation* of the implemented decision allows to control changes and potentially a gain in knowledge for future HIAs.

Quantification takes place within the stage of *effect analysis*, consisting of three different tasks that have to be addressed:

- *Description of the baseline situation:* The current population health has to be described and its future development has to be predicted *in absence* of the proposal under discussion.
- *Estimation of change in exposure to determinants of health:* The causal pathway in HIA assumes that a policy acts on one or more risk factors which in turn affect health. Hence, the change in the risk factor exposure for (parts of) the population caused by the policy has to be estimated.
- *Estimation of change in health outcomes:* The consequences for population health of the new risk factor prevalence are predicted and compared with predictions of the baseline situation, i.e. the situation without an intervention.

Quantification approaches in HIA

Currently, quantification in HIA is attempted seldom (compare Veerman[2] for a detailed overview). A review of 31 guidelines[10] for HIA showed that usually two recommendations for quantification are given: *Use published evidence* or *Commission an expert*. Only few guidelines contained explicit recommendations on quantification. If explicit recommendations are given, then usually for methods in line of an environmental impact assessment or quantitative risk assessment (QRA) framework. Scrutinizing published HIAs and also accounting of research that is in the spirit of HIA, i.e. assessing future health effect of policies, shows the large diversity of approaches used. The approach to quantification taken is often shaped by the particular research questions at hand – scope and data availability – and the research traditions of the field

the conductor of the HIA is most familiar with. Loosely speaking, three approaches dominate the literature.

First, *statistical modeling techniques* – such as multiple regression – are often employed, in particular by economists. The outcome variables are usually one or more harm indicators (e.g. mortality or sick days) at the population level. As explanatory variable some aggregated measure of the exposure variable is used and the effect of changes in the exposure variable on the outcome variable is quantified, often adjusting for other variables. The approach is based on well-established, well understood, and often used statistical/econometrical models. It is relatively modest in data needs as usually only the population level is modeled. However, detailed individual-level epidemiological mechanisms are often ignored. The approach rather assumes that past, aggregate-level relationships hold on to the future. But long time series may not reflect this relationship anymore; short time series may not contain enough information (too few observations to be statistically significant). The choice of variables to adjust for (such as income) can be crucial and needs to be guided by appropriate theory. Furthermore, structural changes could occur in the future which cannot be modeled or foreseen, such as changes in preferences, medical treatment, policy, or population composition.

Second, *quantitative risk assessment* approaches that are the most prominent recommendations in the reviewed HIA guidelines stem from an (epidemiological based) environmental health impact assessment framework.[7,11,12] QRAs explicitly utilize epidemiological evidence such as relative or absolute risks assuming that the exposure in question causes the health outcome. The assessment estimates the population attributable proportion of disease cases or deaths due to the exposure (a family of measures usually described in standard epidemiological texts as PAR, PAF, or sometimes PIF[13,14]) and the expected change in those numbers due to the proposal in question. QRA is much more in line with epidemiological considerations and is able to account for more complex dose-response relationships. But it requires prior research concerning the strength of the causal connection between risk factor and health outcomes. An instructive example is available in the final project report for a health impact assessment for the European Union by Abrahams et. al[15,16] that quantifed the population-level effect of changes in employment policy. This study assessed, inter alia, the effect of a shift from permanent work contracts to fixed-term work contracts on self-reported health using an odds-ratio from published literature.[17,18]

Third, increasingly *dynamic population health models* are being applied to HIA or HIA-related questions[19], such as PREVENT[20], POHEM[21], or ARMADA[22]. In particular ARMADA, developed by University College London (UK), is intended to fit within an environmental/health impact framework. It models real-life populations over time and features a generic disease model with explicit risk factor states. The methodology is rather heterogeneous but usually based on insights of demographic modeling, i.e. life tables. The population is segmented by risk factor status, which in turn determines the probability of transitioning to another state (e.g. death, diseased). This usually requires a larger amount of epidemiological evidence. Thus, models often try to find a way to mitigate this problem, e.g. PREVENT using PIF or POHEM using a tailor-made data base. Dynamic population health models are, in terms of epidemiological and demographic insights, a comparatively accurate way of modeling. However, dynamic population health models are usually not only demanding in their data needs but often require specialist knowledge and/or are only designed for particular risk factors, populations and/or applications.

This thesis

Need for a standard tool

A tool is a device or procedure that can be used repetitively to achieve a set of certain tasks or solve a class of problems. Usually, a single-purpose tool performs comparatively better at its single task than a more general tool that has to be applicable to a much wider range of tasks. Hence, a standard tool intended for a given field must be sufficiently generic to be useful for a large number of problems without neglecting the idiosyncrasy of a particular problem. Nonetheless, a potential standard tool for quantification in HIA has several advantages. First, a standardized approach increases ease of use through repeated experience and, consequently, facilitates routine applications of such a tool within HIA. Second, the use of a standard tool allows comparison of results across applications and research teams. Third, the routine use of a standard tool allows the cumulation of knowledge. And fourth, an established standard tool accepted as valid increases trust in the results. Although the exact specification of what a standard tool for HIA should entail is part of this thesis, at the very minimum such a standard tool for HIA

should be a software that synthesizes epidemiological evidence on risk factor exposure within a population health framework and should be, in principle, applicable across countries, populations and risk factors.

Research questions

This thesis is part of a larger research project, DYNAMO-HIA (see Appendix A for details), and the thesis' contributions are conceptually split into two parts. The first part, *Methodological Foundation*, addresses the technical core in which the requirements for a generic quantification tool for HIA are motivated and specified. Furthermore, the technical choices made in implementing the tool into a self-contained software are described and justified. The second part, *Selected Applications*, presents three applications to explore the adequacy of the tool for the various challenges faced in quantifying health impacts of risk factor modification.

In particular, this thesis aims to answer the following questions:

I. Methodological Foundation

- What requirements must a quantification tool fulfill to be considered a standard tool for health impact assessment?
- Does a tool already exist that fulfills these criteria and could therefore be used as a standard tool for health impact assessment?
- What technical elements are needed for the underlying mathematical model when constructing a standard tool for health impact assessment?

II. Selected Applications

- What are potential health gains and losses for three modifiable lifestyle-related risk factors across Europe?
- What are the EU-wide potential health gains of an established alcohol control policy, i.e. price increase?
- What are the comparative life course health impacts of two complementary, stylized BMI-reducing interventions when utilizing a cohort perspective?

Outline

Chapter 2 conceptualizes the requirements a potential standard tool in HIA should fulfill. These are both, technical and usability criteria ensuring that the two core tasks of an HIA – to predict and to inform – can be sufficiently executed. Furthermore, it demonstrates that as of 2008 no tool existed that could fulfill these criteria. Chapters 3 and 4 describe and outline the technical choices made to satisfy the needs and constraints of the HIA process. The first of these two chapters explains the technical elements and their rationale in more detail, in particular the combination of micro- and macro-simulation, and the reduction in data needs, that are novel for models of such scope. The latter chapter aims to give an accessible overview of the implemented model while having a stronger focus on the adequacy of the model for typical HIA applications.

The following three chapters apply the finished tool to assess its usability for various realistic research questions while exploring potential limitations. Chapter 5 quantifies potential health gains and losses in eleven EU-countries for the risk factors alcohol, BMI, and smoking. To undertake this quantification, "worst"- and "best"-practice risk factor prevalences for three risk factors will be defined and then applied to eleven EU-countries. Chapter 6 applies DYNAMO-HIA to a current policy question, i.e. the increase of the common excise tax on alcohol in the EU, by quantifying the potential health consequences for various price scenarios. Chapter 7 compares the outcomes of two stylized strategies to reduce the levels of obesity using a life course perspective. DYNAMO-HIA is used to compare an intervention that reduces the obesity prevalence when a cohort enters adulthood with an intervention that reduces the lifetime risk of a cohort to become overweight or obese for a cohort.

Chapter 8 summarizes the answers to the research questions succinctly and discusses limitations of the research findings. Additionally, implications for further research are presented. Chapter 9 contains both, an English and Dutsch summary of this thesis. The Appendix contains supplemental information on the data used for the applications. Furthermore, the outline of the DYNAMO-HIA project and the participating collaborators are described.

Part I
Technical Foundation

2

No standard tool for quantification in health impact assessment: A review[†]

Abstract

Background The health impact assessment (HIA) of policy proposals is becoming common practice. HIA represents a broad approach with quantification of the impact of policy options at its core. However, no standard tool is available and it remains unclear whether any current model can serve as a standard for the field.

Purpose The aim of this study is to assess whether already existing models can be used as a standard tool for the quantification step in an HIA.

Method A search in 2008 identified 20 models for HIA of which six are sufficiently generic to allow for *various* and *multiple* diseases and *different* risk factors: ARMADA, Global Burden of Disease, POHEM, PREVENT, Proportional Life Table Method, and RIVM-CDM. These were evaluated along three proposed model structure criteria (real-life population, dynamic projection, explicit risk factor states) and three usability criteria (modest data requirements, rich model output, generally accessible), developed to address the needs and requirements of the HIA framework.

Results Of the six generic models investigated, none fulfills all the proposed criteria as a standard HIA tool. The models are either technically ad-

[†]Stefan K. Lhachimi, Wilma J. Nusselder, Henrik C. Boshuizen, Johan P. Mackenbach. No standard tool for quantification in health impact assessment: A review. *Am J Prev Med*, 38(1):78–84, 2010 (reprinted with permission from Elsevier).

vanced with no or limited accessibility, or they are accessible but oversimplified.

Conclusion Further work on models for HIA with equal emphasis on technical appropriateness, availability of data, and end-user-friendly implementation is warranted if the field is to move forward.

Introduction

Health impact assessment (HIA) assesses the effect of a program, project, or policy on overall population health and the distributional effects within a population.[1] The rationale for HIA is that many risk factors for chronic diseases are affected by policy measures outside the realm of health policy (e.g. transportation, food, or urban planning). Assessments have been carried out at all governmental levels (e.g. local,[23] regional,[18] national,[24] and supranational[15]) and the number of HIAs is likely to rise due to increased institutional adoption and political will, in particular at the European Union level.[25,26]

An HIA can take many forms, ranging from a rapid assessment to establish if health is affected at all, to a comprehensive HIA that appraises all health aspects. Generally, an HIA can be divided into five steps (Figure 2.1). According to a recent definition[2] an HIA consists of two core tasks:

- supporting decision-makers in choosing between options and
- predicting the future consequences of implementing different options.

To predict future developments of complex systems such as a population, models are indispensable.[27] Despite the increasing role of guidelines in HIA and the existence of models that allow to quantify the effect of a policy on population health, there is no commonly accepted practice in the prediction of the impact of policy on health for HIA purposes.[28] Quantification is seldom attempted in HIA[29] and standard tools to conduct this are lacking.[28–31] However, rational decision-making requires that costs and benefits in terms of health effects are estimated.[12,32]

A recently established research consortium (DYNAMO-HIA) aims to develop and make available a standard tool to aid the quantification step in assessing the impact of policy on health by utilizing previously established modeling approaches. A standard tool for quantification in HIA should be

able to quantify the baseline situation (population health without intervention) and then quantify changes due to one or more policy options. Moreover, it must be readily available, and be suitable for repeated HIA exercises and for specific questions. Therefore, the tool should be *generic* in its applicability by allowing for *various* and *multiple* diseases, *different* risk factors, and take into account the standard causal pathway assumed in HIA. The standard HIA causal pathway assumes that a policy intervention leads to a change in risk factor prevalence which, in turn, leads to changes in disease incidence and disease-related mortality and, therefore, in overall population health.[9] Furthermore, it should comply with the needs and objectives of HIA[3] – *to predict the impact* and *inform the decision-making process* – while equally accounting for the constraints of a decision-making process: time and resources (in particular modeling expertise and data) are scarce in an applied setting.

The aim of this study is to assess whether already existing models can be used as a standard tool for the quantification step in an HIA. Consequently, models were included that have been used or are suggested as possible tools for HIA, either by modeling experts or HIA practitioners.

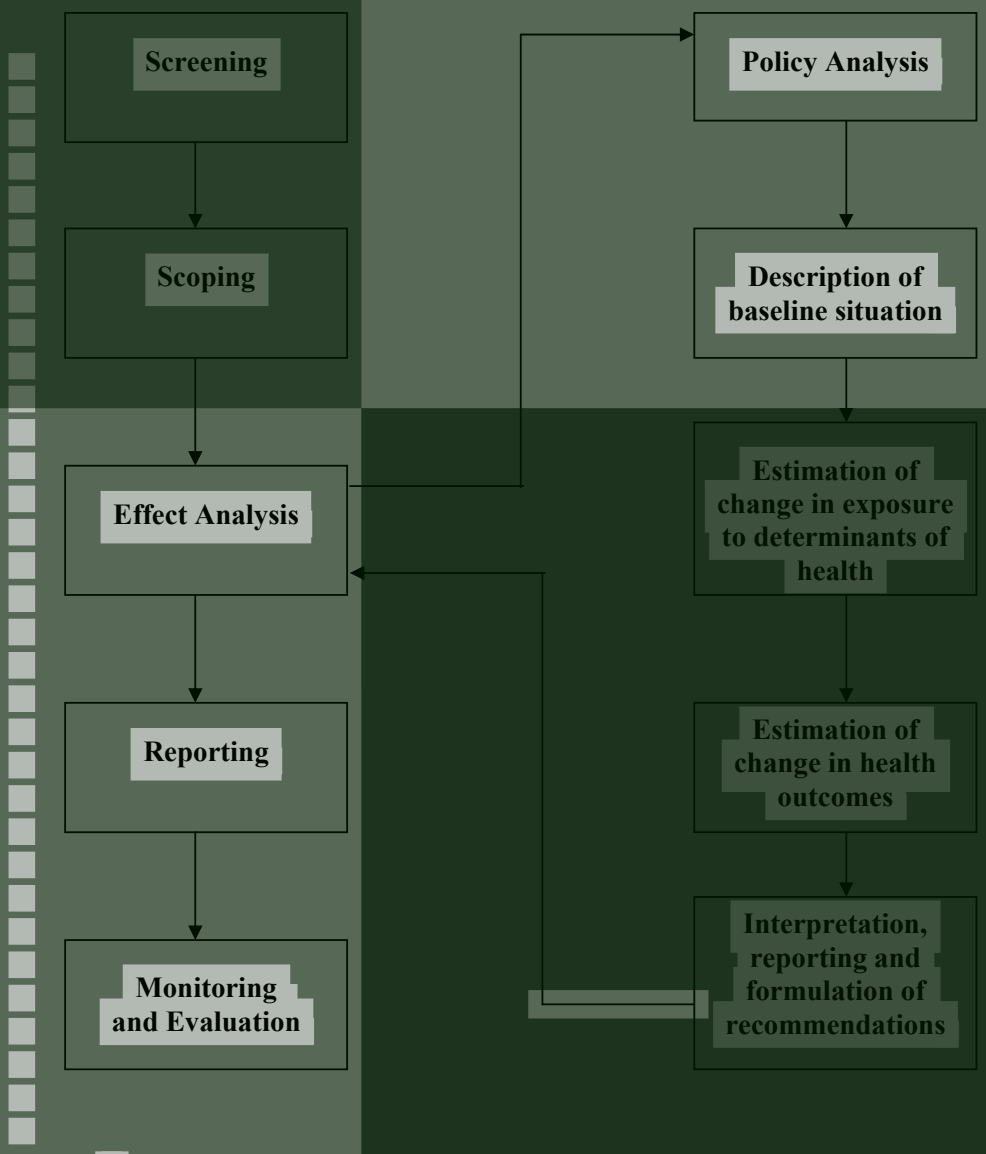

Figure 2.1: Overview of the process of health impact assessment.

Methods

Search strategy

A multi-step procedure was employed (see Figure 2.2 for details). First, a search was made in PubMed (for keywords see Figure 2.2). Second, an informal survey was conducted among modeling and HIA experts, i.e. all members of the consortium of the DYNAMO-HIA project (40 persons from 31 institutions in Europe, Australia and North America), asking whether they knew of any models that are or could be used for HIA. Third, the HIA-dedicated websites `www.who.int/hia` and `www.hiagateway.org.uk` including the links therein were scrutinized: a list of identified models (with a brief outline of the research aim) was then sent to the mailing lists of *HIA-NET* and *HIA-NET South East Asia* to ask for any additional models (both are HIA-dedicated international electronic discussion groups). Fourth, three recent reviews addressing HIA studies were examined for simulation tools: these articles were an exploratory overview of European HIA studies[33], a systematic review of HIAs conducted in the US[34], and a systematic review of quantification in HIA in general[29].

Selection

The present review is restricted to models (e.g. tools that reason and theorize in the language of mathematics to make predictions[27]) that are sufficiently *generic* with regard to the varying subject matters of different HIAs. In the present study, a model is defined as *generic* for the purpose of HIA if it models *multiple* diseases, does not have a pre-set risk factor (e.g. arbitrary risk factors can be simulated), and takes into account the standard causal pathway assumed in HIA. The change in risk factor prevalence due to policy interventions is usually not determined within such a model.

In total, twenty models were identified (see Figure 2.2); of these, six were excluded because they simulate specific diseases and eight because they are risk factor specific (mostly smoking or air pollution) or have no risk factors at all. Thus, six models remained for a comparative review.

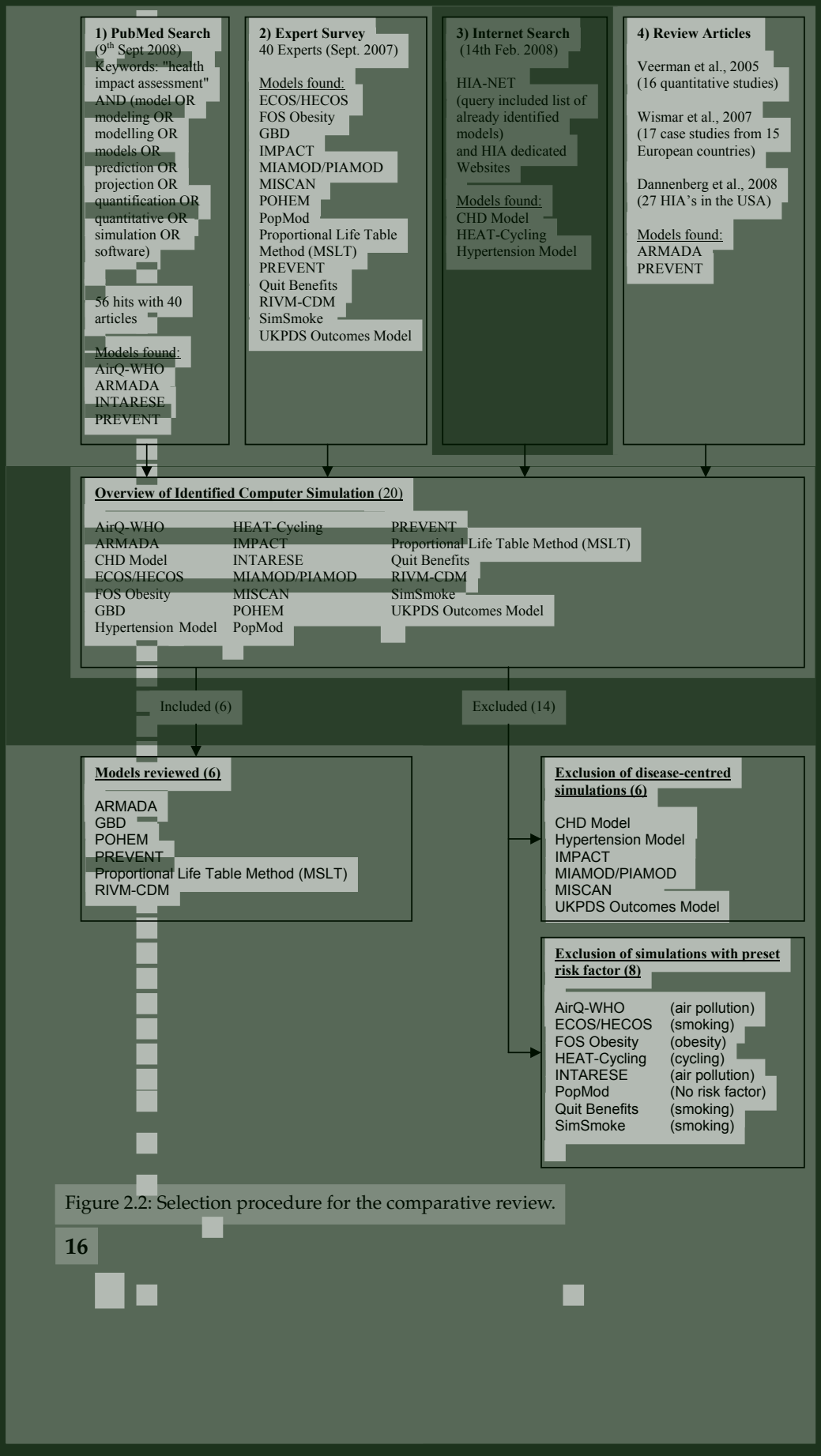

Figure 2.2: Selection procedure for the comparative review.

Evaluation criteria

To evaluate the selected models for their usefulness as a standard tool for HIA, a set of criteria was developed. These criteria were extensively discussed among HIA and modeling experts during the DYNAMO-HIA workshop (May 2008, Rotterdam, the Netherlands).

The two objectives of HIA – to *predict* and *inform* – address the technical core of quantification (*predict*) as well as the context (*inform*) in which HIA takes place. Therefore, a tool should aim for technical accuracy in the prediction of the effects of interventions on population health, and also be effective in the applied setting of an HIA.

We propose six evaluation criteria. The first three criteria (*real-life population*, *dynamic projection*, and *explicit risk factor states*) ensure that the model structure is sufficiently advanced to model changes in risk factor exposure over time in a real-life population in a transparent way. The last three criteria (*modest data requirements*, *rich model output*, and *generally accessible*) ensure a wide usability by accounting for the constraints of a decision-making process.

Real-life population

A tool must be able to model different populations which vary in age- and sex-structure. Age and sex are important confounders for many diseases, and the (re)distribution of health gains and losses within a population is important to fully assess the impact of a policy. A simulation at the patient level, or a simple life table, cannot account for such differences.

Dynamic projection

Changes in population health over time are important for the sustainable assessment of a policy, i.e. apart from information on outcome at a theoretical steady-state, details on how and when it is reached are also needed. Whereas it might be acceptable for a long-term health gain to be achieved by a health decrease in the short term,[19] this needs to be made explicit in an HIA tool.

Explicit risk factor states

For the sake of transparency, it is preferable that risk factor states are explicitly modeled. Therefore, the model should account for the risk factor and

health status of each simulated individual or subgroup at every time period. Models using the potential impact fraction (PIF) approach[35] do not explicitly model risk factor states. This technical simplification ignores mortality selection and may lead to biased estimates.[36]

Modest data requirements

The most commonly available data used to assess the effects of risk factors on population health are population data on *incidence*, *prevalence*, *mortality*, and *relative risk*. The collection of additional data is usually not feasible in an applied setting.[37] In particular, data on incidence, prevalence, and mortality by risk factor status (e.g. for smokers and never-smokers), or on *transitions* between risk factor states, are scarce. Therefore, a model should run on readily available epidemiological data.

Rich model output

HIA informs policy makers about the effect of a policy, and should be done in an impartial manner.[4] A summary population measure might imply a value judgment.[38] On the other hand, simply presenting the number of deaths averted[39] or extensive tables of raw outputs without summary measures, might also be confusing. To enable the end user to choose the appropriate outcome measures, a broad range should be provided (e.g. disease outcomes, number of deaths, *and* summary measures of population health).

Generally accessible

A standard tool should be publicly available and easily accessible. A tool that requires specialized knowledge is accessible to a limited number of experts and therefore less useful. Moreover, there is a significant difference between an internally working model that can be shared with a fellow modeler and a finished piece of software that can be publicly released without further assistance from the developer. The design of the latter requires substantial additional resources to ensure that the software runs in a reliable way.

Results

In the present study, six models were identified that are sufficiently generic:

Table 2.1: Comparison of the reviewed models against the evaluation criteria.

	Criterion					
	Real-life population	Dynamic projection	Explicit risk factor states	Modest data requirements	Rich model output	Generally accessible
ARMADA	+	+	+	-	-	-
GBD	-	-	-	+	-	+*
POHEM	+	+	+	-	+	-
PREVENT	+	+	-	+	+	-
Proportional Multi-state Life Table (MSLT)	-	-	-	+	+	+*
RIVM-CDM	+	+	+	-*	+	-

*with some restrictions

1. ARMADA (Age-Related Morbidity and Death Analysis)
2. GBD (Global Burden of Disease)
3. POHEM (Population Health Modeling)
4. PREVENT
5. MSLT (Proportional Multi-State Life Table)
6. RIVM-CDM (RIVM Chronic Disease Model)

Table 2.1 shows how these six models perform against the evaluation criteria. Because no model fulfills all the criteria, none of the reviewed tools qualifies as a standard tool for HIA. In terms of model structure, the range spans from very complex models that simulate almost all heterogeneity in a real-life population (POHEM), to very simple ones yielding only an approximation of the disease burden of a risk factor (GBD). There are no *a-priori* advantages of a micro- compared to a macro-implementation: POHEM uses a micro approach whereas the majority uses a macro approach, with the exception of RIVM-CDM that includes both. Macro- and micro-implementations yield (in principle) the same results when designed for the same task.[40,41] The former is often easier to implement for less complex population structures, but becomes burdensome when more sophisticated populations are modeled. Therefore, simulations that allow for a high degree of population heterogeneity (e.g. ethnicity, marriage status, risk factor histories, and so on)

opt for a micro-simulation approach. Details on each model are presented below.

ARMADA

ARMADA, developed by University College London (UK), is intended to fit within an environmental/health impact framework (EIA/HIA).[22] It models real-life populations over time and features a generic disease model with explicit risk factor states. A limitation of ARMADA for broader use is that it requires explicit specification of transition rates between every disease- and risk-state combination, thereby increasing data requirements. No summary measures of population health are included. Furthermore, ARMADA is not publicly available and its development is currently suspended. It has been applied to assess the health effects of a planned incinerator plant and traffic accidents.[19]

GBD

The GBD model is a methodology rather than a comprehensive computer simulation and applies the epidemiological effect measure PIF. The model is static, but has modest data needs. Outcome measures are life years lost, disability-adjusted life years (DALY) and healthy life expectancy (HALE). However, the tool has to be implemented by the end user in spreadsheet software. The GBD model has been used in several applications.[42–44]

POHEM

POHEM is a discrete event simulation developed by Michael Wolfson at Statistics Canada. It models the risk factor and health status of individuals during their lifetime that jointly form a real-life population. Tailored to the Canadian context, it uses unique individual (and longitudinal) data on health and other personal characteristics (e.g. socioeconomic status, income, family status). Access to such data is very limited in Europe. POHEM has a broad range of outcome measures and is not publicly available. It is a comprehensive software that has been used in several applications (e.g. cost-effectiveness analysis and policy evaluation for several diseases).[21,45]

PREVENT

PREVENT was developed by Jan Barendregt and models a real-life population over time. It is based on an epidemiological multi-state life table of chronic diseases using PIF to model the effects of interventions on transitions. Consequently, the data requirements are sufficiently modest but mortality selection cannot be handled. Outcome measures include disease outcomes, mortality figures, and summary measures of population health. PREVENT is available from the developer upon request and is not intended for public release. Applications include smoking cessation[20] and increased physical activity[46].

Proportional Multi-State Life Table

The proportional multi-state life table is rather a methodology than a complete model and has to be implemented (e.g. in spreadsheet software) by the end user. It is an extension of a simple life table that incorporates competing causes of death.[47] It models multiple diseases and assesses the effect of interventions using PIFs. Additional weaknesses are lack of dynamic modeling capabilities and that it does not model a real-life population, i.e. depending on the chosen interpretation, it simulates either a single cohort over time or a population in a steady state. However, it is very transparent and uses widely available data. Outcome measures have sufficient breadth (disease outcomes, mortality data, and summary measures such as life expectancy). Examples of applications include the health effects of obesity trends[48] and of changes in agricultural policy[49].

RIVM-CDM

The RIVM-CDM model was developed by the RIVM (National Institute for Public Health and the Environment, the Netherlands) to study the effectiveness of policies for primary prevention and to conduct burden of disease calculations for the Dutch government.[50] It models a real-life population over time and clearly links explicit risk factor states to multiple diseases and death. A feature of the RIVM model is that it uses an additional module to calculate mortality from incidence and prevalence, and estimates transition rates from cross-sectional data, thus reducing data needs; however, the data required are specifically tailored to the Dutch context. RIVM-CDM provides a broad

range of outcome measures, but requires knowledge of a programming language and is not publicly available. Applications include modeling smoking cessation[51] and obesity[52].

Discussion

The present review identified 20 simulation models, of which six were sufficiently generic to possibly serve as a standard tool. All reviewed tools have been used to quantify the effects of changes in risk factor prevalence on population health, but none could comply with the proposed criteria reflecting the needs, objectives, and applied setting of an HIA. With the exception of ARMADA, none of the reviewed models were developed with the intention to be used for HIA. ARMADA was intended for a broad user base, but was not completed due to lack of funding to make the additional changes to accommodate the needs of end users. This demonstrates that the step from an internally working model to finished, publishable software is considerable. For example, the design and implementation of a graphical user interface (generally not required for internal use but essential when used externally) is time and cost intensive.[53]

Similarly, the testing of software is technically demanding and labor intensive. For generally accessible software, testing must be finished prior to publication and cannot be conducted on demand (as is the case with software that is internally developed and used). For example, although PREVENT is available from the developer, it is still undergoing development.[54] Both ARMADA and PREVENT illustrate that considerable time and resources are needed to build a model that not only functions for internal use but is sufficiently developed for a broad user base without requiring repeated contact with its developers.

Furthermore, in the context of HIA, a software model is always a decision-support tool only. It helps to quantify the expected differences in population health given two (or more) different situations: one of them a baseline situation (without the intervention) and one (or more) with intervention(s). It does not predict the development of population health as such. Decision-makers must be constantly aware that real-word phenomena are necessarily more complex and no model can predict future events with perfect accuracy. In HIA it might be useful to avoid calling the results of mathematical models "predictions" but rather *projections* of "what-if" scenarios in a clearly de-

fined and simplifying framework. Mathematical models allow to synthesize current knowledge with clearly-stated assumptions in a logical fashion and, consequently, can yield useful insights for decision making. The term "prediction" could be reserved for the entire process, in which a software model is only one element of the evidence base.

Study limitations

In all systematic reviews, publication bias is a notable problem because "grey" or non-English literature is often not included in standard databases.[55] Models that could be useful for HIA, but have not been publicly used, are therefore not covered. In the present study, this limitation was decreased by using a structured, multi-step approach in the search procedure. The survey among modelers and HIA practitioners was particularly successful in revealing models not mentioned in the HIA-based literature; e.g. the PubMed search of HIA-related literature revealed only two models (ARMADA and PREVENT).

Assessing the *validity* of the reviewed models is beyond the scope of this study. Although suggestions are available in the health field, no distinct standard emerges.[56,57] Moreover, it is difficult to evaluate validity without having full access to the software and extensive experience with the individual models.

Conclusion

This study shows that no existing model can serve as a standard tool for quantification in HIA. There is an evident gap between the advanced models that have no or limited general accessibility (such as POHEM and RIVM-CDM), and the (over-)simplifying but generally accessible models (such as GBD and MSLT). This situation probably arises because none of the reviewed models (except for ARMADA) was initially intended to be a software application for wide public use for the (relatively recent) task of quantification in HIA. In addition, the laudable, but (as yet) partly realized plans for ARMADA show that significant resources and a strong focus on usability are needed to make a model accessible to a wider audience.

This review illustrates that any tool that intends to fill this gap in the future needs to put equal weight on an appropriate simulation methodol-

ogy and data requirements that can be widely met while also being end-user friendly. Consequently, even at the design phase, the DYNAMO-HIA consortium attempted to take into account the needs and capabilities of the targeted end user and put aside sufficient resources to make an internally working model ready for public use, whilst also minimizing data needs.[58]

Acknowledgments

The authors are grateful to the members of the DYNAMO-HIA project and consortium for discussing the evaluation criteria and for participating in the survey of potential tools for quantification in HIA. The authors also thank members of the HIA-Net and SE-HIA-Net for their participation in the survey.

3

The DYNAMO-HIA model: An efficient implementation of a risk factor/chronic disease Markov model for use in health impact assessment[†]

Abstract

In health impact assessment (HIA) or priority setting for health policy, effects of risk factors (exposures) on health need to be modeled, e.g. with a Markov model, in which exposure influences mortality and disease incidence rates. As many risk factors are related to a variety of chronic diseases these Markov models potentially contain a large number of states (risk factor and disease combinations), providing a challenge, both technically (keeping execution time and memory use down) and practically (estimating the model parameters and retaining transparency).

We propose a combination of micro-simulation of exposure information with macro-simulation of diseases and survival to handle this task. This allows simulation of exposure detail while avoiding the need for large simulated populations due to the relative rareness of chronic disease events. Further efficiency is gained by splitting the disease state space into smaller spaces each containing a cluster of diseases that is independent of the other clusters. The challenge of feasible input data requirements is met by including parameter calculation routines which use marginal population data to estimate both the transitions between states and the initial state occupancy.

[†]Henrik C. Boshuizen, Stefan K. Lhachimi, Pieter van Baal, Rudolf T. Hoogenveen, Jet Smit, Johan P. Mackenbach, Wilma J. Nusselder (accepted for publication in Demography)

As an illustration, we present the recently developed DYNAMO-HIA model implementing this approach.

Introduction

In medical demography, demographic techniques are employed to model the consequences of morbidity, disability and mortality for the size, composition, and structure of the population. Especially, medical demography is concerned with how changes in the incidence and prevalence of specific diseases impact the patterns of other co-morbid conditions, disability and mortality. This can be extended by modeling these changes in incidence or prevalence as consequences of changes of underlying determinants or risk factors. Such models have found applications in the field of health impact assessment (HIA)[11,23,59] and in priority setting for health policy[60,61] as they offer a way to model effects of interventions or policies on population health. Such modeling proceeds by assuming that policies/interventions change the prevalence of particular risk factor states in the population (for instance the proportion of smokers), or change the transition rates between risk factor states (for instance, the rate of starting or stopping smoking). Demographic population health models then project the effects of the new risk factor situation by assuming that the risk factor status influences disease incidence and mortality rates (Figure 3.1). The effects of the risk factor change (and thus of the policy/intervention) then can be calculated by comparing the projected size, composition (especially in terms of health state), and structure of the population under these new conditions with that modeled under *business-as-usual* conditions.

For this type of demographic models, a Markov model is a natural choice. In such models each person is characterized by a current state (presence or absence of modeled diseases and risk factor level). The model projects the future prevalence of states (state occupancy) by repeatedly applying a matrix of transition probabilities to the vector of current state occupancy. Such a matrix of transition probabilities contains the probabilities of changing from each current state to each possible next state.

In real-life applications, such Markov models often need a large number of states, as risk factors such as smoking and obesity are related to multiple chronic diseases. States are formed by all possible combinations of disease and risk factor states, and thus increase exponentially with the number of

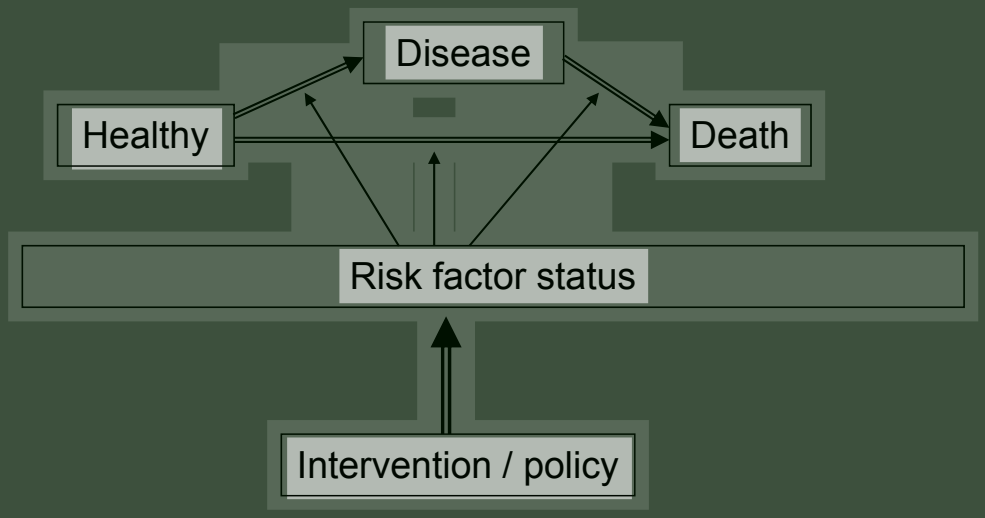

Figure 3.1: General idea of a model for health impact assessment or priority setting in health policy: A policy or intervention will influence the risk factor status of individuals, which in turn influences their probabilities of becoming diseased or dying.

diseases and risk factors included. Also, the central assumption of a Markov model (the Markov property) is that the transition probabilities only depend on the current state, and not on previous states. In practice, however, the probability of dying from a disease, for instance, can depend on duration of exposure or time since diagnosis. This can be handled by extending the number of states (in the example: let state include the aspect *exposure time* or *time since diagnosis*), resulting in a model satisfying the Markov property, but having more states. Furthermore, many risk factors (like BMI or alcohol consumption) are measured on a continuous scale, and most easily modeled in the form of a large number of distinct values, also increasing the number of states. For use in applied settings, as HIA or priority setting, the choice of a particular simulation methodology is constrained by time and resources (in particular, modeling expertise and data).[62] Hence, designing a model with a large number of states is disadvantageous in several ways. First, the model should be fast enough for routine use. Second, with many states, the number of transitions between states gets extremely large, and getting observational data on all transition rates individually is not feasible. Therefore, a system is needed to derive transition rates from more general data. Lastly, despite increased complexity, transparency should be maximized. Current modeling either opts for using in-house, computer intensive solutions based on micro-

simulation[63,64], or circumvents these problems by modeling only population average exposure[46,65], using approximations[66] or limiting the scope of the model[62]. In this chapter, we propose a set of solutions that can be used in demographic population health models when a relatively simple and fast running program is needed for use in applied settings. The first aspect of our approach is that it combines micro-simulation of risk factor (exposure) information with macro-simulation of diseases and survival. This allows simulation of sufficient risk factor detail (e.g. continuous risk factor, time since exposure started or stopped) while avoiding the need for large simulated populations. Second, we present an efficient way to split the disease state space into smaller state spaces, each for an independent cluster of diseases, which can be updated and inspected separately. This increases not only the speed of the model but also increases transparency as disease clusters can be inspected separately. Third, introduction of random noise in the micro-simulation part is restricted to a minimum by applying variance reduction techniques. We will illustrate our solutions with a practical application, the DYNAMO-HIA model, in which these solutions are implemented, including a description on how data requirements where kept to a minimum. Our solutions made it possible to build software that can include multiple diseases and detailed information on risk factor exposure, can be filled with available data and runs with reasonable speed for routine calculations.

Methods for computation

Macro-simulation

In a Markov model, a matrix of transition probabilities is repeatedly applied to a vector of state probabilities (state occupancies). A deterministic or *macro-simulation* approach simply carries out such a repeated application of the transition matrix. For a model with N states, this needs a N$\times$N transition probability matrix. When the number of states becomes large, either due to a large number of diseases or due to use of many different risk factor states, this approach quickly becomes infeasible. For example, a model of the effects of smoking including 12 smoking states (never, current, and 10 classes of former smokers by duration of stopping), 15 smoking-related diseases, 90 categories for age, and 2 categories for gender, has $12\times2^{15}\times90\times2\approx70$ million states, and a transition matrix with 5×10^{15} entries. Approximate methods have been de-

veloped to handle such Markov models with many states,[66] but they require some strict assumptions.

Micro-simulation

Another solution is to implement the Markov model as a stochastic or *micro-simulation* model. Instead of keeping track of the expected probability (occupancy) of every single state, this approach acknowledges that many of these states are very rare (e.g. having 4 particular diseases at the same time). Micro-simulation takes a representative sample of *individuals* from the state-space, where the probability that an individual with a particular state is in the sample is proportional to the probability of that state. The sample could contain multiple individuals in the same state (when this state is common), while very rare states are often not included in the sample. Then the life courses of these individuals are simulated from the transition probabilities: Transition rates are used to draw (using a random number generator) a next state for each simulated person (fixed time-step simulation), or to draw a waiting time to the next change of state (discrete event simulation).

Such a model implements the same Markov model, only the manner of calculation is different: The macro-simulation model calculates the probability of every possible state exactly (wasting time on extremely rare states that do not impact the overall result), while the micro-simulation model simulates the distribution over the states in a sample of persons. When simulating a sufficiently large sample, the proportion of simulated persons in a particular state will represent the probability of that state.

The partial micro-simulation approach

The macro-simulation approach for a model with N states needs the complete $N \times N$ transition rate matrix, while the micro-simulation approach needs the states of the M simulated persons, and M transition rate vectors of size N at each moment in time. Therefore, micro-simulation will require less computational resources than macro-simulation when $M < N$. However, the random drawing of states in micro-simulation will add random variation that needs to be averaged out by using a sufficiently large M.

In our design, we acknowledged that using a micro-simulation approach seems most suitable for modeling risk factors, as risk factors are characterized by a large number of states, either because they are continuous, or be-

cause of aspects as time since starting or stopping exposure. On the other hand, a macro-simulation approach is more suitable for modeling diseases, as incidence rates of many diseases in the general population are low, so large samples are needed for stable results. Therefore, we propose a combination of both approaches: *partial micro-simulation*.

Partial micro-simulation implies that for each simulated individual, the risk factor history is generated similarly to a fixed time-step micro-simulation. At each moment in time each individual has a *risk factor state*, and this state is updated by randomly drawing a new risk factor state using the transition matrix between risk factor states. In contrast, the disease part uses macro-simulation. Instead of assigning an individual to either the *with disease* state or the *without disease* state, the individual is assigned a probability of having the disease. Similarly, death is not simulated by assigning *death* or *alive*, but by assigning a probability of being alive.

In case of multiple diseases, the *disease probability* consists of a series of probabilities, one for each combination of all separate disease states. For instance, in case of two diseases A and B there are 4 disease probabilities: the probability of having neither disease, the probability of having only disease A, the probability of having only disease B, and the probability of having both diseases.

For those unaccustomed to working with probabilities, one can also see this as taking a sample of risk factor histories, and then calculating a multi-state life table[67] separately for each risk factor history. An essential condition for using this approach is that the risk factor transition rates are not influenced by the disease states of an individual. For example, the change in BMI can not depend on the presence or absence of diseases. In our application, where incidence of diseases depends only on risk factor status before getting the disease, this restriction is justified.

After running the simulation, one can average the survival, or the probability of a particular disease state, over all simulated individuals, or over a subgroup of simulated individuals (only males, or only 40 year olds), yielding the average probability of the disease state or survival in this group. Also, at this stage (DALY) weights can be attached to the disease states, which can similarly be averaged over groups, enabling the calculation of disability-adjusted life expectancy (DALE).

Separate handling of disease clusters

With a large number of diseases, the number of disease combinations can still be quite large. Calculation of multi-state life tables with a large number of states is time consuming when this needs to be repeated many times, as in partial micro-simulation. However, as shown in Technical Supplement, if – conditional on the risk factor history – transition rates for a disease (or for a cluster of diseases) can be taken to be independent of the transition rates for another disease (or cluster of diseases), then the multi-state life table can be split into a series of independent tables: one per disease (or cluster of diseases). Each, now smaller, table can then be updated independently. This reduces simulation time and also increases transparency, as instead of one large table with disease state probabilities, there is a series of smaller, more manageable tables. When a cluster contains only a single disease, this *table* only contains the probability of having the disease. This approach was already proposed by Barendregt[47] for single diseases, but can be extended to clusters of diseases (see Technical Supplement A for proof).

Variance reduction methods

We employ three well-established variance reduction methods to further mitigate the random noise of the micro-simulation approach.

First, models for health impact assessment and priority setting for health policy aim to compare interventions or scenarios. The variance in this comparison, i.e. between two or more scenarios, can be reduced by keeping the random components of the risk factor exposure equal for the same individual under different scenarios/interventions. This ensures that the random component of the risk factor simulation affects the differences between scenarios only in second order.[68]

Second, additional variance is introduced when an initial population is generated using a random generator to draw a risk factor state for each individual. To minimize this variance, the initial risk factor distribution is not randomly drawn, but assigned deterministically, making the risk factor distribution as equal as possible to the targeted risk factor distribution.

Third, for categorical risk factors, any discrepancies between the targeted risk factor distribution and the realized distribution in the simulated population is dealt with by weighting the results, in order to deliver results that at the start of simulation have exactly the targeted risk factor distribution.

Table 3.1: Input of the DYNAMO-HIA model.

For simulation	For post-processing (scaling to population numbers, and calculation of summary measures of population health)
The (marginal) distribution of the risk factor in the population (e.g. the percentage of smokers, or mean BMI with standard deviation)	Population Numbers
the (marginal) probability of having a particular disease (e.g. the prevalence of heart disease)	Projected numbers of newborns
the incidence rate of a particular disease (e.g. the incidence of heart disease)	Daly weights for each disease
The relative risk from the risk factor on each disease	Daly weights for the entire population
The all-cause mortality rate	
The excess mortality rate from each disease	
Transition rates for the risk factor	
Optionally: The relative risk of a disease on getting another disease (independent of the effect of the risk factor)	
Optionally: The relative risk of the risk factor on all-cause mortality	

The DYNAMO-HIA model

As an illustration, we present the recently developed DYNAMO-HIA model, a model developed with health impact assessment in mind, but which can also be used for priority setting in the health policy context. DYNAMO-HIA first projects the population under *business-as-usual* conditions, and alternatively under one or more scenarios in which the risk factor distribution in the population has been changed. The model as implemented only allows for a single, but generic risk factor in the model, which can be freely chosen by the user. However, the *single* risk factor could be a combination of risk factors (as: drinking smokers, non-drinking smokers, drinking non-smokers and non-drinking, non-smokers). As this example shows, this limitation is a feature of the current implementation, amendable in the future.

In DYNAMO-HIA, the methods described above were implemented. Furthermore, DYNAMO-HIA applies an epidemiological model that defines the

Table 3.2: Model parameters to be supplied to the simulation part of the DYNAMO-HIA model.

Transition to	Parameter	Description
Another risk factor states	$\lambda_{i \to j}(\Delta t = 1)$	1-year transition probability from risk factor category i to risk factor category j (categorical risk factor)
	$\partial(\Delta t = 1)$	Annual change in risk factor level (continuous risk factor)
	$\sigma(\Delta t = 1)$	Size of random term that is annually added to the continuous risk factor in order to increase the variance with age
Diseased	$I_{0,i}$	The baseline incidence of disease *i*, that is the incidence in the state where all relative risks are equal to 1
	$RR_{r \to d_i}$	The relative risk of risk factor state *r* on disease *i*
	$RR_{c_j \to d_i}$	The relative risk of disease *j* on disease *i*
Death	Am_i	Attributable mortality, that is the mortality due exclusively to disease *i*
	$RR_{r \to OC}$	Relative risk of risk state r on other cause mortality (other then from the diseases in the model)
	$M_{0,OC}$	Baseline other cause mortality
	M_{0,CF_i}	Baseline fatal incidence rate of disease *i*
	$RR_{r \to d_i} RR_{c_j \to d_i}$	Relative risks for acutely fatal disease, identical to the relative risks for incidence

transition rates using only a limited set of parameters, as well as a module that generates an initial population from marginal population data (Technical Supplement B). After running a simulation on a simulated sample of the initial population, in a post-processing step the simulation results are scaled up towards the population numbers of the real population, and summary measures of population health, such as life expectancy or disease-free life expectancy, are calculated. In this step also Disability Weights (DALY-weights) can be attached to disease states in order to calculate disability-adjusted life years (DALE; see Technical Supplement C for further details).

Epidemiological model in DYNAMO-HIA

For running a risk factor/disease Markov model, one needs a matrix of transition rates (e.g. how many smokers with diabetes die, or become smokers with diabetes and heart disease). In practice it is not feasible nor necessary for a user interested in doing an HIA to obtain all these data at this level of detail. Instead of working with an unrestricted matrix of transition rates, transitions can be parameterized applying an epidemiological model that defines the transition rates based on a limited set of parameters. This epidemiological model defines transition rates in continuous time, i.e. transitions take place in an infinitely small period of time. Technical Supplement D describes how these transition rates are converted into transitions over larger timesteps during the simulation. The epidemiological model of DYNAMO-HIA parameterizes two transition rates: incidence rates and mortality rates. Remission rates are not included but the model could be extended to include them.

Incidence rates

The transition rate from not having disease *i* to having disease *i* is called the (non-fatal) incidence rate of *i*, and is described by the following equation:

$$I_i(r, \mathbf{C}) = I_{0,i} RR_{r \to d_i} \prod_{c_j = c_1}^{c_m} RR_{c_j \to d_i} \tag{3.1}$$

where

r	is a risk factor state
$\mathbf{C}$	is a vector of states of the causal diseases, that is, diseases that are a cause of another disease with elements $\{c_1,\ldots, c_m\}$, e.g. for 5 causal diseases: {0,0,1,0,0}
$I_{0,i}$	is the baseline incidence of disease i, that is, the incidence in the state where all relative risks are equal to 1
$RR_{r\rightarrow d_i}$	is the relative risk of risk factor state r on disease i
$RR_{c_j\rightarrow d_i}$	is the relative risk of causal disease j on disease i

In this equation the baseline incidence $I_{0,i}$ is estimated from the input data as described in Technical Supplement C, while the relative risks are input to the program.

Mortality rate

DYNAMO-HIA contains three different disease processes determining mortality:

1. a chronic disease process, which is characterized by a constant (attributable) mortality rate after getting the disease that only depends on age and gender
2. a disease process in which the disease can be acutely fatal
3. a disease process including a cured fraction

Diseases with constant mortality rate

For these diseases, the transition rate to death given risk factor state r and disease state $\mathbf{D}$ (the mortality rate M(r,$\mathbf{D}$)) is given by:

$$M(r,\mathbf{D}) = M_{0,OC} RR_{r\rightarrow OC} + \sum_{i=1}(Am_i d_i) \tag{3.2}$$

where

$M(r, \mathbf{D})$	The mortality rate given risk factor state r and disease state **D**
r	is the risk factor state
D	is a vector of states of diseases with elements $\{d_1, \ldots, d_n\}$
Am_i	is a model parameter referred to as attributable mortality, giving the mortality rate due exclusively to disease i
$RR_{r \to OC}$	is the relative risk of risk state r on other cause mortality (that is, other than from the diseases in the model)
$M_{0,OC}$	is the baseline other cause mortality

The dependence of the other cause mortality on risk factor status is optional in the model. Without this dependence the other cause mortality is the same for all simulated subjects, and the risk factor has only an indirect effect on mortality, that is only through its effect on disease incidence.

Diseases that can be acutely fatal

For these diseases, a term for acutely fatal diseases is added to the mortality rate $M(r, \mathbf{D})$ as given by (3.2):

$$M(r, \mathbf{D}) = M_{0,OC} RR_{r \to OC} + \sum_{i=1} Am_i + \sum_i CF_i(r, \mathbf{D}) \tag{3.3}$$

where

$CF_i(r, \mathbf{D})$	$= M_{0,CF_i} RR_{r \to d_i} \prod_{c_j=c_1}^{c_m} RR_{c_j \to d_i}$
c_j	is an element of the vector of causal disease states **C**
M_{0,CF_i}	is the baseline fatality rate of disease i
$RR_{r \to d_i}$	is the relative risk of risk factor state r for disease i
$RR_{c_j \to d_i}$	is the relative risk of causal disease j for disease i

This mortality rate is meant for diseases where the mortality in the first period after disease onset is much higher than in later periods. The model assumes that the fatal cases die at or shortly after disease onset, while the non-

fatal cases are subjected to a constant attributable mortality over time. Those with non-fatal disease are still at risk of getting fatal disease. This disease process was introduced with cardiovascular diseases in mind, where fatal strokes and fatal myocardial infarctions commonly occur in those with earlier strokes and myocardial infarctions respectively.

Diseases with a cured fraction

A disease with a cured fraction is basically split up at time of diagnosis in two diseases: a *cured disease* and a *not cured disease*. The latter have an increased mortality rate due to the disease (attributable mortality), which is constant over time, the former has an attributable mortality of zero. The disease process was introduced with cancer in mind. It accommodates a mortality pattern where the attributable mortality slowly declines towards zero with time since diagnosis.

For both the cured and non-cured disease $M(r, \mathbf{D})$ is given by Equation 3.2 setting Am_i to zero for the cured disease. The excess mortality of the disease (user input to the model) is used to calculate the attributable mortality of the not-cured disease, while the cured disease gets a zero attributable mortality. Cured cases are no longer at risk for getting the disease.

Model input needed

Some of the parameters needed for this epidemiological model, namely the relative risks on diseases, should be directly given by the user of the model. Other parameters are estimated by the DYNAMO-HIA parameter-estimation module from epidemiological population data (as percentage of smokers and incidence of heart disease). These methods are largely taken from similar procedures used in the RIVM-CDM[66] and are described in Technical Supplement B, together with the way in which an initial population is estimated from these data.

Example

As an example, we calculate the hypothetical scenario in which the present (year 2004) average BMI in the Dutch population is reduced to the values of 1990. As the distribution of BMI is slightly skewed (Figure 3.2), we modeled BMI with a log-normal distribution. After preliminary model-selection, we

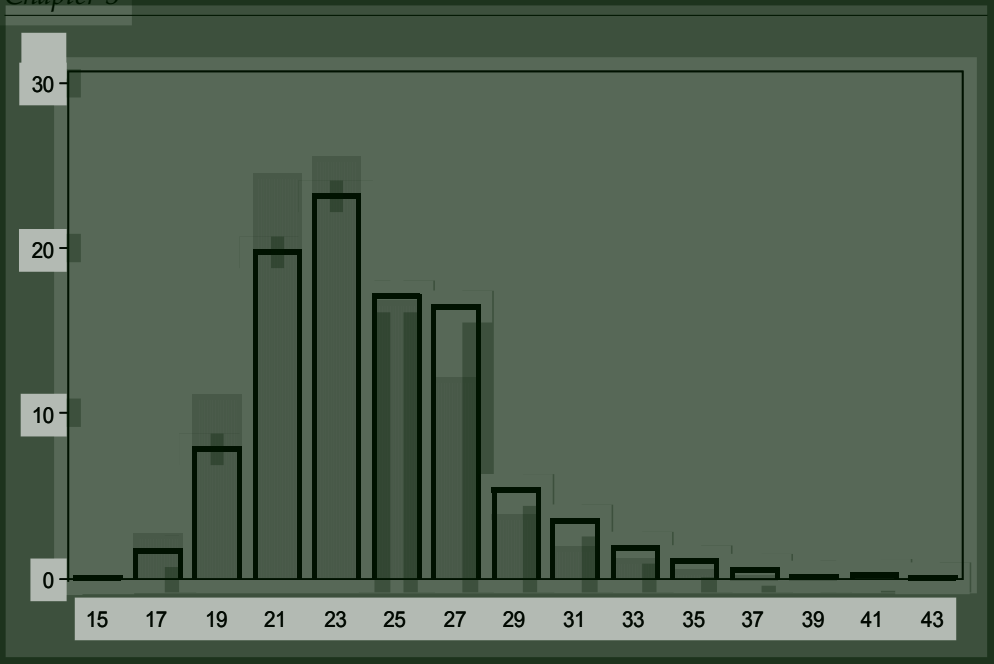

Figure 3.2: Distribution of BMI (x-axis) in 1990 (grey) and 2004 (black lines) in percent (y-axis). Source: Health Interview Survey 1990 and Periodic Survey of the Life Situation 2004 (POLS) of Statistics Netherlands.

fitted a third grade polynomial model of age with log(BMI) as dependent variable, separately for both genders, including an interaction term with year and assuming constant residual error variance on the log(BMI) scale (Figure 3.3).

Transitions rates for BMI were chosen so that the average age-specific BMI remained constant over time (thus equal to the 2004 and 1990 distributions, respectively; see Technical Supplement B for details). Diseases in the model were ischemic heart disease, stroke, diabetes, colorectal-, esophageal-, and breast-cancer. Diabetes was considered to be a causal disease for ischemic heart disease and stroke. Stroke and ischemic heart disease were modeled as diseases that were acutely fatal (see section on epidemiological model), and no relative risk on total mortality was used. We used disease data and relative risks from the DYNAMO-HIA tutorial data set.

Figure 3.4 and 3.5 illustrate the parameter calculation procedures used. Figure 3.4 shows the input *incidence of diabetes in the entire population* together with the parameter *baseline incidence of diabetes*, that is, the diabetes incidence in those with a BMI of 22.5 kg/m^2. It shows that while around age 70 the diabetes incidence in the input data is higher in women than in men, this is

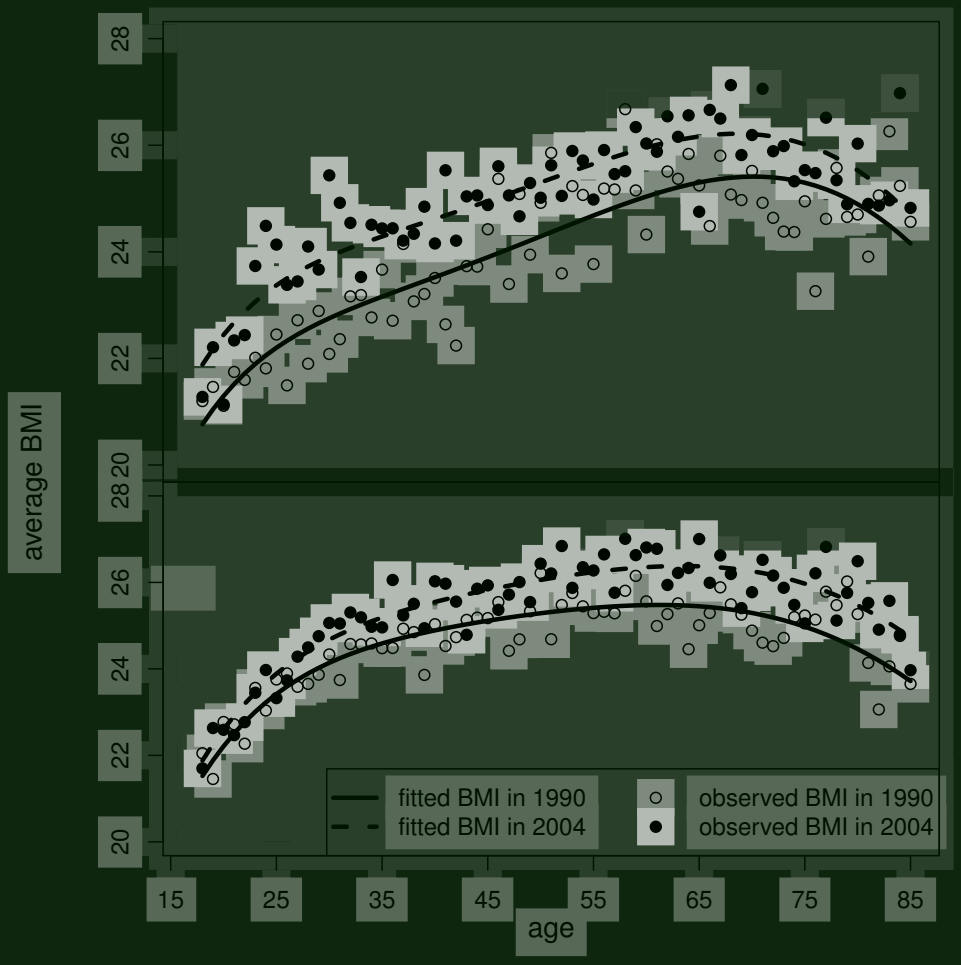

Figure 3.3: Fitted average BMI values by age for men (lower graph) and women (upper graph) on the data from the Health Interview Survey 1990 and Periodic Survey of the Life Situation 2004 (POLS) of Statistics Netherlands.

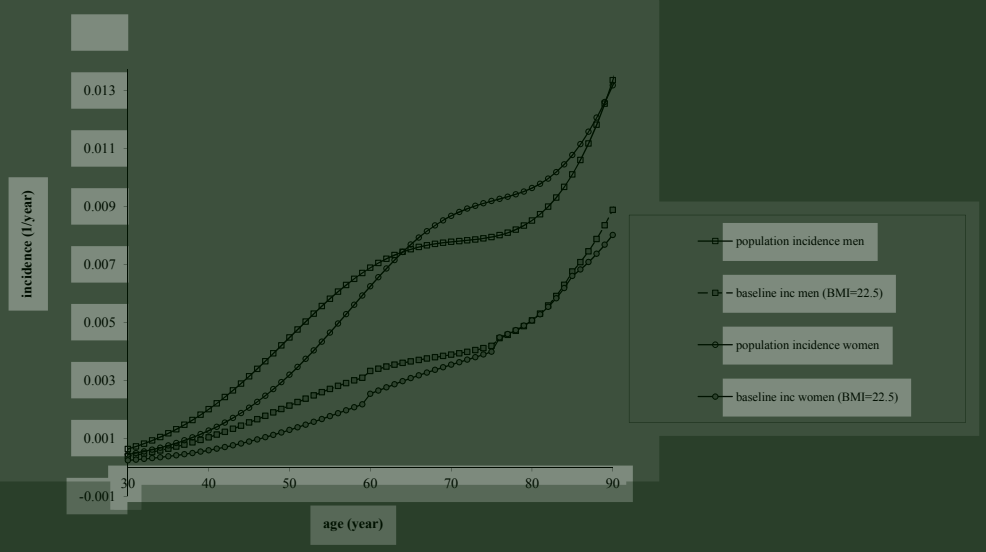

Figure 3.4: Comparison of the incidence rate of diabetes (input of to model) and the parameter that is calculated from this input (the baseline incidence, in this case the incidence in individuals with a BMI of 22.5.

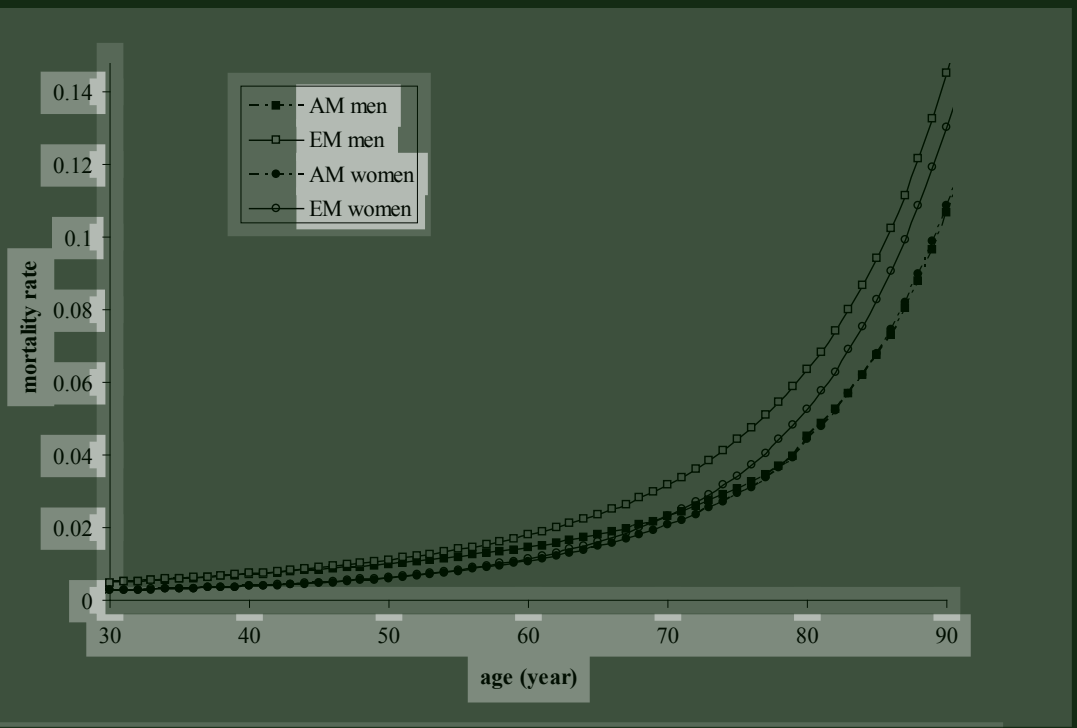

Figure 3.5: Comparison of the excess mortality (EM) for diabetes (input of the model) and the parameter that is calculated from this input, the attributable mortality (AM) for diabetes.

no longer the case for the baseline incidence: The higher population incidence of diabetes in women of this age results from their on average higher BMI. The slight discontinuities, visible in the baseline incidence around the ages 60 and 75, are due to the discontinuity of the relative risks used, which were constant in broad age groups (e.g. a single relative risk was used for all ages 60-75).

Figure 3.5 shows the input excess mortality for diabetes, as well as the attributable mortality calculated from it. This calculation removes from the excess mortality the mortality due to the higher prevalence of ischemic heart diseases (IHD) and stroke in diabetes patients compared to the general population. At young ages, the effect of this exclusion is small, but at higher ages the attributable mortality is clearly lower than the excess mortality.

Figure 3.6 shows a result screen of the DYNAMO-HIA model that summarizes the simulated data by displaying a population pyramid at simulation year 20. In this pyramid, the number of persons that have at least one of the 6 diseases in the model is shown in pink. The number of diseased persons (that is, persons having at least one of the modeled diseases) in excess of that of the reference scenario (2004) is given in red, while the reduction in the number of diseased persons relative to the reference scenario is given in yellow. Similarly, the reduction in the number of persons that are alive is given in grey, while any excess is given in black. In this example, we see that reducing the BMI in the population to the 1990 levels at most ages leads both to less mortality (larger population) and a lower number of diseased persons in the population. A very close look, shows that at the highest ages the number of diseased persons is larger under the *return-to-1990 BMI-values* scenario (that is, lower BMI). This is because more persons stay alive under the alternative scenario, and part of those have a disease. From the simulated data, one can calculate that the cohort life expectancy of a newborn changes from 77.66 under the *business-as-usual* scenario to 77.84 under the return-to-1990 values scenario in men, and from 82.04 to 82.27 years in women, while the disease-free life expectancy changes from 70.46 to 70.90 in men and from 73.79 to 74.40 in women. As indicated by Figure 3.6, other plots can be made in the DYNAMO-HIA user interface to study different aspects of the simulated data. In addition, an option is provided to output the simulated numbers for further processing in other programs.

Discussion

We described an approach, implemented in the DYNAMO-HIA model, combining the best of existing model approaches in order to efficiently program a Markov-model based disease model. We combined micro-simulation for risk factor modeling with macro-simulation for calculating disease prevalence in order to minimize running time for a particular accuracy.

This approach, although still flexible, is more limited than a full micro-simulation approach. In partial micro-simulation a simulated individual is not assigned a disease state but only a disease probability. Therefore, modeling situations where the transition rates depend on time since disease onset or where risk factor changes depend on the disease state are problematic. For mortality, the first restriction is alleviated in the DYNAMO-HIA model by providing two disease processes (*with cured fraction* and *acutely fatal*) that model a) the case where excess mortality due to a disease declines exponentially with time since diagnosis, and b) the case where there is a short window of high mortality immediately after diagnosis, respectively. For other situations, where transition rates depend on time since diagnosis, full micro-simulation might be more appropriate.

The risk factor histories, however, are simulated using micro-simulation, and here transitions can depend on the length of time in a particular risk factor state. In the current DYNAMO-HIA software this is only implemented for a single risk factor class, but conceptually this can easily be extended.

In partial micro-simulation, each combination of diseases forms a different state. We avoid the curse of dimensionality by splitting the disease-space into disease clusters that are mutually independent, given the risk factor. This assumes that such independent clusters exist. Although in theory many dependencies between diseases could exist, in practice many are weak (e.g. cancer seems to be largely independent of that of cardiovascular diseases, conditional on joint risk factors[67]) and/or data to quantify the amount of dependence is lacking. Hence, modeling them as independent is the most reasonable option. In order to explore the consequence of unjustly assuming independence, we also carried out our example calculation without the correlation between diabetes, stroke, and IHD. The years with disease changed by less than 4%, and the difference in disease-free years between scenarios by only 1%. As the dependency between diabetes, stroke, and IHD is stronger than the dependency between most other diseases, this indicates that the as-

sumption of independence will not unduly influence results.

A Markov model projects both future disease states (disease prevalence), and future transition rates (incidence and mortality). Projecting future prevalence rates of disease is relatively rare in chronic disease modeling: many models only project incidence and mortality. Projecting prevalence, however, is necessary in order to calculate summary measures of population health like health expectancy or DALEs that are useful both in the HIA context and for priority setting. The reason for the scarcity of models that project prevalence into the future is that such a projection requires mutually consistent prevalence, incidence and excess mortality data in order to prevent unrealistic projections. Checking data quality and consistency is thus a prerequisite for models projecting prevalence. The DisMod II software[69] has been developed to check such consistency. Although requiring consistency complicates the process of getting input data for the model, this requirement is not particular to our model: for example, methods for calculating DALYs require disease duration as input, which is based on equivalent DisMod calculations.

Most of the data used in modeling are subject to uncertainty. Probabilistic (or Monte-Carlo) uncertainty analysis[70–72] can be used to estimate the uncertainty in model outcomes from the uncertainty in the data. To facilitate use of such methods, the DYNAMO-HIA model can be run in batch mode, making it possible to automate rerunning it many times on data sets generated in such a Monte-Carlo approach.

In summary, we believe our approach is a good compromise between flexibility, reasonably short running times and realistic data needs. Therefore, the approach was implemented in the DYNAMO-HIA model for use in health impact assessment.

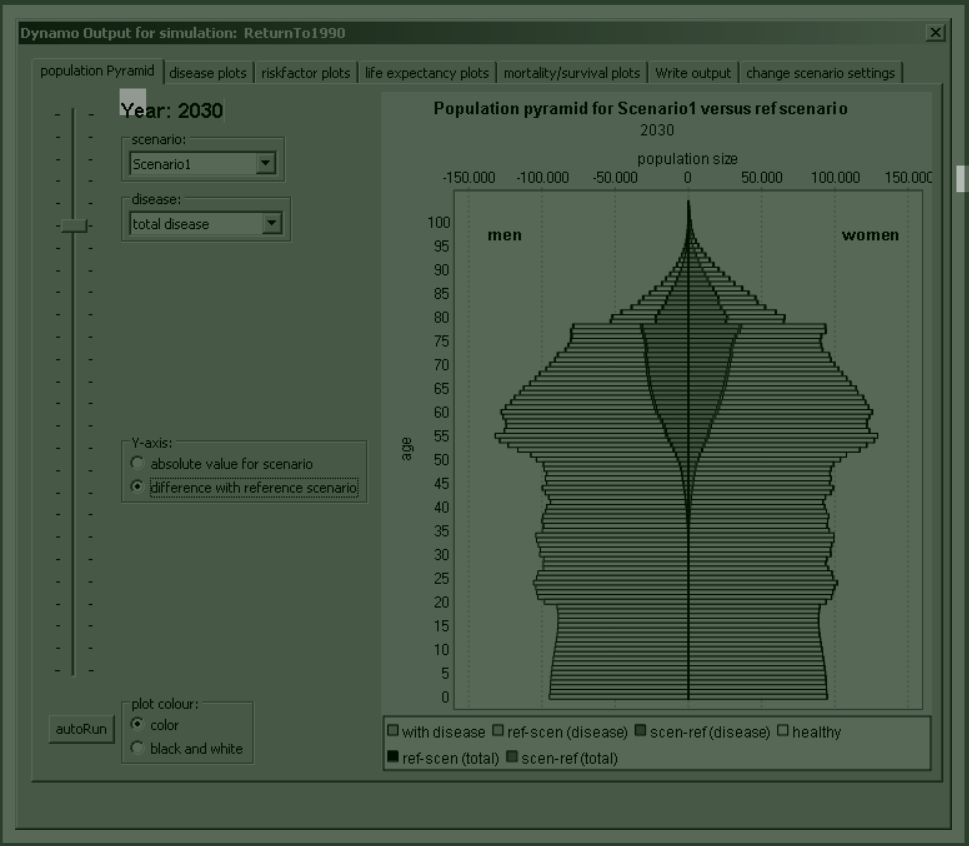

Figure 3.6: DYNAMO-HIA output: the population pyramid after simulating 20 years under the alternative *return-to-1990* scenario.

Technical Supplement A: Partitioning the calculation of the transition probability matrix for clusters of independent diseases

In the proposed model, there is for every simulated individual a transition state probability matrix $\mathbf{\Pi}(\Delta t)$ that describes the transition rates between all disease states for a time-step Δt. This matrix is given by:

$$\mathbf{\Pi}(\Delta t) = e^{\mathbf{Q}\Delta t} \tag{3.4}$$

where $\mathbf{Q}$ is the matrix of transition rates. $\mathbf{Q}$ has a dimension $(2^n+1)\times(2^n+1)$, where 2^n is the number of distinct disease states in the model (n being the number of diseases), and death is the last, absorbing, state.

$\mathbf{Q}$ is of the form:

$$\mathbf{Q} = \begin{bmatrix} q_{11} & q_{12} & q_{13} & \cdots & q_{1(2^n+1)} \\ q_{21} & q_{22} & q_{23} & \cdots & q_{2(2^n+1)} \\ q_{31} & q_{32} & q_{33} & \cdots & q_{3(2^n+1)} \\ \vdots & \vdots & \vdots & & \vdots \\ q_{2^n 1} & q_{2^n 2} & q_{2^n 3} & \cdots & q_{2^n(2^n+1)} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

with $\sum_j q_{ij} = 0$ and $q_{ii} \leq 0$ and $q_{ij} \geq 0$ for $i \neq j$

To construct this matrix, only the $2^n \times 2^n$ upper left matrix, further referred to as $\mathbf{T}$, is needed, as the remaining entries can be derived using $\sum_j q_{ij} = 0$.

Given the definition of the matrix exponential,

$$e^{\mathbf{Q}} = I + \sum_{k=1}^{k=\infty} \frac{1}{k!}\mathbf{Q}^k = \mathbf{I} + \mathbf{Q} + \frac{1}{2}\mathbf{Q}^2 + \frac{1}{6}\mathbf{Q}^3 + \cdots$$

it is clear that $\mathbf{\Pi}(\Delta t)$ will be of the form

$$\mathbf{\Pi}(\Delta t) = \begin{bmatrix} \pi_{11} & \pi_{12} & \pi_{13} & \dots & \pi_{1(2^n+1)} \\ \pi_{21} & \pi_{22} & \pi_{23} & \dots & \pi_{2(2^n+1)} \\ \pi_{31} & \pi_{32} & \pi_{33} & \dots & \pi_{3(2^n+1)} \\ \vdots & \vdots & \vdots & & \vdots \\ \pi_{2^n 1} & \pi_{2^n 2} & \pi_{2^n 3} & \dots & \pi_{2^n(2^n+1)} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

with $\sum_j \pi_{ij} = 1$, as the last row of $\mathbf{Q}^2$ (and thus also other powers of **Q)** will only contain zeros and its other rows will sum to 0. Entries in the last column of **Q** will only influence the entries in the last column of $\mathbf{Q}^2$ (and other powers of **Q)**, and thus the upper left $2^n \times 2^n$ matrix of $\mathbf{\Pi}(\Delta t)$, further referred to as $S(\Delta t)$, is equal to the matrix exponential of $\mathbf{T}\Delta t$. Given $S(\Delta t)$ and the constraints $\sum_j \pi_{ij} = 1$, the matrix $\mathbf{\Pi}(\Delta t)$ can be derived and, thus, the problem of finding $\mathbf{\Pi}(\Delta t)$ from **Q** can be simplified to finding $\mathbf{S}(\Delta t)$ from **T**.

As a first step, calculation of $\mathbf{S}(\Delta t)$ is partitioned into separate calculation of 1) the effect of *other cause mortality* (OC), that is mortality that does not depend on the particular disease state, and 2) disease-related transitions. As mortality rates only cause transitions out of the current state (into death, which is not part of the states in **T**), mortality rates will contribute only to the diagonal terms of **T**. As OC does not depend on the disease state, all diagonal terms of **T** contain the same term – OC, and the transition rate matrix **T** thus can be written as $\mathbf{T}^* - OC\mathbf{I}$, where **I** is the $2^n \times 2^n$ identity matrix, and **T*** the transition rate matrix without the OC terms, containing only terms from incidence, recovery and disease attributable mortality rates. For commuting matrices **A** and **B** it holds that $e^{\mathbf{A}+\mathbf{B}} = e^{\mathbf{A}} + e^{\mathbf{B}}$. It can easily be seen that **T*** and – OC **I** are commuting, and thus

$$e^{\mathbf{T}} = e^{-OC\mathbf{I}} e^{\mathbf{T}*} = e^{-OC} e^{\mathbf{T}*}$$

therefore partitioning the calculation of $\mathbf{S}(\Delta t)$ in calculating $e^{\mathbf{T}*}$ and e^{-OC} separately.

As a next step, we will show that calculation of e^{T*} can further be partitioned in multiplying separate matrices for mutually independent clusters of diseases. We will show this below for two clusters of diseases. Given that

this is true for two clusters of diseases, it is clear that it is also true for any arbitrary number of diseases: One can first combine the matrices from the first two disease clusters if the property holds for two disease clusters. This yields a matrix for the two clusters together (a new first cluster), and this then can be combined with the third cluster, and so on.

Let there be two independent clusters of diseases: cluster 1, containing the set of disease states C with elements c_k, and cluster 2, containing the set of disease states D with elements d_m. The overall set of states comprises all combinations $c_k d_m$. Independence of the clusters of diseases means that:

1. The transition rate between state of $c_k d_m$ and $c_l d_m$ is the same for all m, and similarly the transition between $c_k d_m$ and $c_k d_n$ is the same for all k. In other words, incidence, recovery and attributable mortality of diseases in a cluster do not depend on the presence or absence of a disease or disease-combination in the other cluster.

2. The prevalences of the diseases are independent of that of the diseases in the other cluster, i.e. (conditional on being alive) the probability of being in state c_k of cluster 1 and in state d_m of cluster 2 is given by:

$$\Pr(c_k \cap d_m) = \Pr(c_k) \Pr(d_m)$$

Note that as the transition rates all apply to a single individual, the requirement of independence only implies independence conditional on risk factor status.

Furthermore, as we are dealing with rates, all transitions should take place in an infinitely small time period. Therefore, only a single transition is possible within this period, so it is not possible to acquire more than one disease, and all transitions between states that differ by the presence of two (or more) diseases are zero. This implies that for both $k \neq l$ and $m \neq n$ the transition between $c_k d_m$ and $c_l d_n$ is zero.

We will illustrate our proof with an example where we let cluster 1 be arbitrary, and cluster 2 contain two diseases. From this example, however, we will generalize to many diseases in cluster 2.

The first cluster, $\mathbf{A}_1$, is the matrix of the transition rates between the states c_k, where the number of diseases is N and their mutual dependence is not further specified. The second contains two diseases, the states d_m are: 00 (both diseases absent), 01 (disease 1 present), 10 (disease 2 present), and 11 (both diseases present). The matrix $\mathbf{A}_2$ of the transition rates between the states

within cluster 2 is then of the form:

$$\begin{bmatrix} -\lambda_{11} & \lambda_{12} & \lambda_{13} & 0 \\ \lambda_{21} & -\lambda_{22} & 0 & \lambda_{24} \\ \lambda_{31} & 0 & -\lambda_{33} & \lambda_{34} \\ 0 & \lambda_{42} & \lambda_{43} & -\lambda_{44} \end{bmatrix}$$

For illustration we also give this matrix, replacing λ_{ij} with the corresponding epidemiological parameter:

$$\begin{bmatrix} -inc_1 - inc_2 & rec_1 & rec_2 & 0 \\ -inc_1 & -inc_3 - am_1 - rec_1 & 0 & rec_4 \\ -inc_2 & 0 & -inc_4 - am_2 - rec_2 & rec_3 \\ 0 & -inc_3 & -inc_4 & -am_3 - rec_3 - rec_4 \end{bmatrix}$$

where

inc_1: incidence rate of disease 1 in those without disease 2
inc_2: incidence rate of disease 2 in those without disease 1
inc_3: incidence rate of disease 1 in those with disease 2
inc_4: incidence rate of disease 2 in those with disease 1
rec_1: recovery rate of disease 1 in those without disease 2
rec_2: recovery rate of disease 2 in those without disease 1
rec_3: recovery rate of disease 1 in those with disease 2
rec_4: recovery rate of disease 2 in those with disease 1
am_1: attributable mortality in those with disease 1
am_2: attributable mortality in those with disease 2
am_3: attributable mortality in those with both disease 1 and 2

The overall matrix of transition rates (between all disease states combining both clusters of diseases) is an N×4 matrix, which we can arrange in blocks with similar states of cluster 2 diseases. As transition rates can only be non-zero when there is only a single change between states, cluster 1 transitions are zero in all off diagonal blocks (which already imply a change in cluster 2 diseases), and the block will consist of $\lambda_{ij}\mathbf{I}$, as λ_{ij} is the same for all states c_k in this block due to the independence between the clusters. Similarly, the off-diagonal elements of the diagonal blocks will not contain any cluster 2 transitions while the diagonals of these blocks will be equal to the corresponding entries of A_1. Only on the diagonal of the matrix transitions are possible concerning both clusters, comprising the diagonal elements of A_1 and $\lambda_{ii}\mathbf{I}$. So the resulting matrix can be written as:

$$\mathbf{T} = \begin{bmatrix} A_1 - \lambda_{11}\mathbf{I} & \lambda_{12}\mathbf{I} & \lambda_{13}\mathbf{I} & 0 \\ \lambda_{21} & A_1 - \lambda_{22} & 0 & \lambda_{24} \\ \lambda_{31}\mathbf{I} & 0 & A_1 - \lambda_{33}\mathbf{I} & \lambda_{34}\mathbf{I} \\ 0 & \lambda_{42}\mathbf{I} & \lambda_{43}\mathbf{I} & A_1 - \lambda_{44}\mathbf{I} \end{bmatrix}$$

This matrix can be written as the sum of 2 matrices, each containing only transition rates from one of the two clusters:

$$\begin{bmatrix} -\lambda_{11}\mathbf{I} & \lambda_{12}\mathbf{I} & \lambda_{13}\mathbf{I} & 0 \\ \lambda_{21} & -\lambda_{22} & 0 & \lambda_{24} \\ \lambda_{31}\mathbf{I} & 0 & -\lambda_{33}\mathbf{I} & \lambda_{34}\mathbf{I} \\ 0 & \lambda_{42}\mathbf{I} & \lambda_{43}\mathbf{I} & -\lambda_{44}\mathbf{I} \end{bmatrix} + \begin{bmatrix} A_1 & 0 & 0 & 0 \\ 0 & A_1 & 0 & 0 \\ 0 & 0 & A_1 & 0 \\ 0 & 0 & 0 & A_1 \end{bmatrix}$$

More generally, we can write for two transition matrices $\mathbf{A}_1$ and $\mathbf{A}_2$ with dimensions N and M respectively:

$$\mathbf{T} = \mathbf{A_2} \otimes \mathbf{I_N} + \mathbf{I_M} \otimes \mathbf{A_1} = \mathbf{A_2} \oplus \mathbf{A_1}$$

where $\oplus$ stands for the Kronecker sum, and thus

$$e^{\mathbf{T}} = e^{\mathbf{A_2} \oplus \mathbf{A_1}} = e^{\mathbf{A_2}} \otimes e^{\mathbf{A_1}}$$

Next, we will prove, that when independent clusters of diseases have prevalence rates that are mutually independent at the start of the simulation (t_0), the prevalence rates will remain independent during simulation. With independence of prevalence we mean that:

$$\Pr(c_k \cap d_m | alive) = \Pr(c_k | alive) \Pr(d_m | alive)$$

where c_k and d_m are particular disease states within disease cluster 1 and 2, respectively.

We will now use the symbol $\mathbf{p}_1(t)$ here for the vector of probabilities $\Pr(c_k|alive)$ (the vector of prevalence rates of cluster 1 disease states) and similarly $\mathbf{p}(t)$ for the vector $\Pr(c_k \cap d_m|alive)$).

At the start of the simulation, all subjects are alive, and the vector of state probabilities $\mathbf{P}(t_1)$ is equal to $\mathbf{p}(t_0)$ and can thus be written as:

$$\mathbf{P}(t_0) = \mathbf{P}_2(t_0) \otimes \mathbf{P}_1(t_0)$$

where $\mathbf{P}_1(t_0)$ is the vector of state c_k probabilities for the first cluster of disease, and $\mathbf{P}_2(t_0)$ is the vector of state d_m probabilities for the second cluster of disease.

The probability $\mathbf{P}(t_1)$ with $t_1=t_0+\Delta t$ is then given by:

$$\begin{aligned}\mathbf{P}(t_1) &= e^{\mathbf{T}\Delta t}\mathbf{P}(t_0)\\ &= e^{-OC\Delta t}(e^{\mathbf{A}_2\Delta t} \otimes e^{\mathbf{A}_1\Delta t})(\mathbf{P}_2(t_0) \otimes \mathbf{P}_1(t_0))\\ &= e^{-OC\Delta t}(e^{\mathbf{A}_2\Delta t}\mathbf{P}_2(t_0)) \otimes (e^{\mathbf{A}_1\Delta t}\mathbf{P}_1(t_0))\\ &= e^{-OC\Delta t}\mathbf{P}_2(t_1) \otimes \mathbf{P}_1(t_1)\end{aligned}$$

where $\mathbf{P}_2(t_1)$ is introduced as shorthand for $e^{\mathbf{A}_2\Delta t}\mathbf{P}_2(t_0)$ and can be interpreted as the probabilities of states in the second cluster of diseases at $t=t_1$ in the hypothetical case that only mortality from this cluster would occur. To calculate $\mathbf{p}(t_1)$ we need to divide $\mathbf{P}(t_1)$ by the survival at time t_1, which is equal to $e^{-OC\Delta t} \times surv_1 \times surv_2$ with $surv_x = \sum_k p_{x,k}(t_1)$, where $p_{x,k}(t_1)$ is the k^{th} element of vector $\mathbf{P}_x(t_1)$.

In thus calculating $surv_1$ and $surv_2$, we use the fact that the state *death* is not part of $\mathbf{P}$, so that the columns of the transition matrices $\mathbf{A}_1$ and $\mathbf{A}_2$ do not add up to 1 but to something smaller, implying that with every update the size of $\mathbf{P}_1(t)$ and $\mathbf{P}_2(t)$ will diminish due to death related to the cluster diseases.

Dividing $\mathbf{P}(t_1)$ by the survival at time t_1 yields

$$\mathbf{p}(t_1) = \frac{e^{-OC\Delta t}}{e^{-OC\Delta t}}\frac{\mathbf{P}_2(t_1)}{surv_2} \otimes \frac{\mathbf{P}_1(t_1)}{surv_1} = \mathbf{p}_2(t_1) \otimes \mathbf{p}_1(t_1) \tag{3.5}$$

Therefore, in order to run the model, one does not need to store all $N \times M$ disease states in $\mathbf{P}(t)$, but it is sufficient to store the disease states $\mathbf{p}_1(t)$ and $\mathbf{p}_2(t)$ (N+M states), and additionally keep track of the overall survival. Similarly, one does not need to exponentiate the (N×M) × (N×M) disease state matrix $\mathbf{T}$, but can update $\mathbf{p}_1(t)$ and $\mathbf{p}_2(t)$ independently with the smaller matrices $\mathbf{A}_1$ (N×N) and $\mathbf{A}_2$ (M×M).

Technical Supplement B: Estimation of initial state occupancy and model parameters from marginal data

In this section, we describe how the DYNAMO-HIA model uses marginal data on risk factor and disease prevalence, incidence and mortality to estimate

1. the initial occupancy of states, and
2. the parameters for the transition rates between states.

The methods used are largely taken from similar methods used in the RIVM-CDM.[66] All calculations as described below are performed separately by age and gender.

Initial occupancy of states

The central assumption used for converting marginal distributions of the individual components of a state (that is the distribution of the risk factor and of diseases) into a joint distribution is that the prevalence odds ratio of having a disease, conditional on all other aspects of the state (that is risk factor status or the presence/absence of other diseases), is equal to the relative risk of this disease. The justification for this is that under certain conditions the ratio of the prevalence odds of a disease equals the ratio of incidences.[73,74] The assumption is inspired by a similar assumption used in the RIVM-CDM,[66] where, however, this assumption is made for the ratio of the prevalence.

In the case of diseases that only depend on the risk factor state r (we further refer to those as independent diseases), we solve the baseline odds of disease i iteratively from:

$$\Pr(d_i = 1) = \sum_r \frac{RR_{r \to d_i} * Baseline\ odds_i}{RR_{r \to d_i} * Baseline\ odds_i + 1} \Pr(R = r) \quad (3.6)$$

where

Pr(d_i=1)	(marginal) prevalence of disease *i* (input to the model)
Pr(*R*=*r*)	(marginal) probability of the risk factor state *r* (input to the model)
$RR_{r \to d_i}$	relative risk of risk factor state *r* on disease *i* (input to the model)

With these baseline odds we then can calculate the joint prevalence of risk factor states and independent diseases. If we then define a new risk factor state r* as the joint risk-factor independent disease state, we can repeat the procedure for diseases that depend on other diseases (referred to as dependent diseases) using r*. Although in theory one could proceed like this, adding extra layers of new dependent diseases in each step, the current implementation of the DYNAMO-HIA model is restricted to a single layer of dependent diseases.

The joint distribution of all risk factor and disease states is used to initialize the simulated population. For this, first, risk factor states are assigned until the number of simulated persons assigned almost (or completely) reaches the expected number under the intended distribution. Second, the remaining simulated persons are randomly assigned a risk factor value based on the difference between the realized and the targeted risk factor distribution. Given the assigned risk factor, the disease state is then calculated based on Equation 3.6 (including the relation between independent and dependent disease).

Transition Rates

Table 3.1 gives an overview of the input needed in the DYNAMO-HIA model, and Table 3.2 of the parameters that govern the transition rates in the DYNAMO-HIA. The parameters $RR_{r \to d_i}$ and $RR_{c_j \to d_i}$ need to be given directly by the user, while $\lambda_{i \to j}(\Delta t)$ and $\partial(\Delta t)$ can either be given directly by the user or can be calculated by the program such that the age-specific prevalence stays constant over time (so called *net-transition* rates). All other parameters need to be estimated from the input (Table 3.1).

Transitions between risk factor states ($\lambda_{i \to j}(\Delta t = 1)$, $\partial(\Delta t = 1)$, $\sigma(\Delta t = 1)$)

Transition probabilities for risk factors can either be given directly by the user, or the program can estimate the minimal transitions needed in order to let the

age-specific prevalence stay constant over time (net-transitions). For categorical risk factors this method was described before.[75] For normally distributed continuous risk factors, the risk factor level L at age $a + \Delta t$ is given by:

$$L(a + \Delta t) = L(a) + \partial(\Delta t) + \sigma(\Delta t)\varepsilon,$$

where ε is a random number with distribution $N(0, 1)$.

For net-transitions, $\partial(\Delta t = 1)$ is simply the difference of the average risk factor level L in the population at age *a+1* and age *a*: $\partial(\Delta t = 1) = \overline{L(a+1)} - \overline{L(a)}$. Similarly, $\sigma(\Delta t = 1)$ is estimated from the increase in variance in the population with age. If the variance decreases with age, $\sigma(\Delta t = 1)$ is made zero; $\sigma(\Delta t = 1)$ is always calculated by the program, also when the user supplies $\partial(\Delta t = 1)$. Relative risks on all-cause mortality are used to adjust these parameters for selective mortality before calculating net-transition rates.

Estimation of the baseline incidence ($\mathbf{I_{0,i}}$)

As the population incidence is the average of the state-specific incidence, using the joint distribution of states in the initial population (as estimated above), we can estimate $I_{0,i}$ from the marginal incidence rate as:

$$I_{0,i} = \frac{I_i(1 - FF_1)}{\sum_r \sum_{C \in \chi} RR_{r \to d} \Pr(C, R = r | d_i = 0) \prod_{i:c_i=1} RR_{c_i \to d}} \tag{3.7}$$

where

FF_i	fatal fraction (user supplied)
$\mathbf{C}$	is a vector of states of the causal diseases (e.g. for 5 causal diseases: [0,0,1,0,0]), with elements {$c_1, \ldots, c_m$}
χ	the set of all possible vectors $\mathbf{C}$ (all possible combinations of presence/absence of causal diseases)
I_i	the (marginal) incidence rate of (non-fatal) disease *i* (user-supplied)

The overall incidence I_i and the fatal fraction FF_i are directly supplied by the user. For diseases that are not directly fatal, the fatal fraction is zero.

Estimation of the attributable mortality ($\mathbf{Am_i}$), baseline other cause mortality ($\mathbf{M_{0,OC}}$), and relative risks for other cause mortality ($\mathbf{RR_{r\to OC}}$)

The mortality rate, given risk factor state r and disease states $\mathbf{D}$, is given by:

$$M(r,\mathbf{D}) = \sum_{i:d_i=1} Am_i + M_{0,OC} RR_{r\to OC} + \sum_i CF_i(r,\mathbf{D}) \tag{3.8}$$

where $CF_i(r,\mathbf{D}) = M_{0,CF_i} RR_{r\to d_i} \prod_{j:c_j=1} RR_{c_j\to d_i}$. Here we have multiple unknowns: The attributable mortalities Am_i for each disease, the relative risks $RR_{r\to OC}$ for risk factor value r on other cause mortality, and the baseline other cause mortality $M_{0,\overline{OC}}$. The marginal population data available for estimating these parameters are the all-cause mortality *M*, the prevalence of disease *i* ($\Pr(d_i = 1)$), the relative risks of each risk factor state on all-cause mortality $RR_{r\to M}$ and the excess mortality rates E_i.

The excess mortality for disease i is defined as the observed difference of mortality in the group with the disease and the group without the disease. This is not equal to the attributable mortality Am_i, as excess mortality includes mortality that results from more co-morbidity and higher risk factor exposure in those with the disease, while attributable mortality does not include those effects. In many situations, however, the difference is small.

To estimate these parameters, first $RR_{r\to M}$ and M are used to calculate *M(R=r)*, the all-cause mortality in risk factor state r. Second, for each disease the excess mortality E_i, the prevalence of disease *i* and the total all-cause mortality M are used to calculate the all-cause mortality in those with disease *i*, $M(d_i = 1)$:

$$M(d_i = 1) = M + (1 - \Pr(d_i = 1))E_i$$

Third, we write down equations for *M(d_i=1)* and *M(R=r)* in terms of the parameters that have to be estimated:

$$M(d_i = 1) = Am_i + \sum_{j \neq i} Am_j \Pr(d_j = 1 | d_i = 1) + M_{0,OC} \sum_r RR_{r \to OC} \Pr(R = r | d_i = 1) + \sum_j CF_j(d_i = 1) \tag{3.9}$$

$$M(r) = \sum_j Am_j \Pr(d_j = 1 | R = r) + M_{0,OC} RR_{r \to OC} + \sum_j CF_j(r) \tag{3.10}$$

where

$$CF_j(d_i = 1) = M_{0,CF_j} \sum_{C \in \chi} \sum_r \Pr(\chi = C, R = r | d_i = 1) RR_{r \to d_i} \times \prod_{k:c_k=1} RR_{c_k \to d_j} CF_j(R = r)$$
$$= M_{0,CF_j} \sum_{C \in \chi} \Pr(\chi = C | R = r) RR_{r \to d_j} \prod_{k:c_k=1} RR_{c_k \to d_j}$$

and

$CF_j(d_i = 1)$	incidence of acutely fatal disease *j*, given that disease *i* is present
$CF_i(R = r)$	incidence of acutely fatal disease *i*, given risk factor state *r*
$\mathbf{C}$	vector of causal disease states with elements c_k
χ	the set of all possible vectors $\mathbf{C}$ (all possible combinations of causal disease states)

Using the joint distribution of all risk factor and disease states as estimated before when defining the initial occupancy of states, all conditional probabilities – such as $\Pr(d_j = 1 | d_i = 1)$, $\Pr(d_i = 1 | R = r)$, $P(\chi = C | R = r)$ and $P(\chi = C, R = r | d_i = 1)$ – are known. Given these known terms, with L diseases and N risk factor states we then have a linear system of L+N equations with L+N unknowns (L *Am*-terms, N-1 $RR_{r \to OC}$-terms, as $RR_{r \to OC}$ is set to 1 for the reference group, and $M_{0,OC}$), which can be solved using stan-

dard linear algebra.

Using $RR_{r\to OC}$ is optional: if this option is not used, risk factors only influence mortality through causing more disease. In that case, there are only L+1 equations to solve:

$$M(d_i = 1) = Am_i + \sum_{j \neq i} Am_j \Pr(d_j = 1 | d_i = 1) + M_{0,OC} + \sum_j CF_j(d_i = 1)$$

$$M = \sum_i Am_i P(d_i = 1) + M_{0,OC} + \sum_i \sum_r CF_i(R = r) \Pr(R = r)$$

Estimation of the baseline fatal incidence rate ($\mathbf{M_{0,CF_i}}$)

The baseline incident $M_{0,CF}$ can be estimated by:

$$M_{0,CF_i} = \frac{FF_i I_i}{\sum_r \sum_{C \in \chi} RR_{r\to d_i} \Pr(R = r, \chi = \mathbf{C}) \prod_{k:c_k=1} RR_{c_k \to d_i}} \quad (3.11)$$

where FF_i is the fatal fraction of disease i (user input). Again, the joint distribution of all risk factor and disease states, as estimated before when defining the initial occupancy of states, is used to supply $P(\chi = \mathbf{C}, R = r | d_i = 1)$.

Technical Supplement C: Definition of DALY weights in DYNAMO-HIA

The DALY-weights are defined as the percentage of decrease of the value of life due to the disease. The DALY-weight Q attached to a particular disease state $\mathbf{D}$ is calculated as:

$$Q = 1 - (1 - Q_0) \prod_{i:d_i=1} (1 - Q_i) \quad (3.12)$$

where: Q_0 DALY-weight for a person without any of the diseases in the model; Q_i DALY-weight due to the disease (supplied by user)

The DALY-weight Q_0 for a person without any diseases is calculated from the overall DALY-weight for the entire population Q as:

$$Q_0 = 1 - \frac{(1-Q)}{\sum_{D \in \xi} P(\xi = \mathbf{D}) \prod_{i:d_i=1} (1 - Q_i)} \quad (3.13)$$

where: ξ Set of all possible disease states $\mathbf{D}$; d_i state of disease i in $\mathbf{D}$

Again, all these calculations are carried out separately for age and gender and $P(\xi = \mathbf{D})$ is estimated in the same way as described in the section on defining the initial occupancy of states.

Technical Supplement D: Justification of the numerical methods used

Fixed time-step simulation implies that a process in continuous time is simulated in discrete time, and thus transition probabilities (from the states at the beginning of the time-step to the states at the end of the time-step) have to be calculated from the continuous time transition rates. Several approaches are being used for this in disease models.

Many models, e.g. the RIVM-CDM[66] use the Euler forward method, basically ignoring the difference between rates and probabilities. If time-steps or rates are small or if the uncertainty about the rates (due to data uncertainty) is much larger than the error from this approximation, this is a defendable approach. Others[76] use more sophisticated numerical methods of forward integration. Analytical solutions to Equation 3.2 are available when transition rates $\mathbf{Q}(\Delta t)$are assumed constant within the time-step Δt, as the matrix of transition probabilities $\mathbf{\Pi}(\Delta t)$ in that case is equal to the matrix exponential of $\Delta t \mathbf{Q}(\Delta t)$ (see Technical Supplement A). However, for all but the most simple cases it is more convenient to compute the matrix exponential using numerical methods, as done in the ARMADA model.[22] Lastly, if transitions are in one direction only, constant within the time-step, and depend only on

the disease state at the beginning of the time-step, one can calculate the probability of going from state A to state B somewhere during the time-step as $1 - e^{-\alpha t}$, where α is the transition rate from A to B. For transitions to absorbing states these probabilities are the probabilities that are needed. For other states the probabilities can be calculated from these probabilities using probability calculus. This method is often used in multi-state life table models and is used as far as we understand it in the UKPDS model.[77]

The first two methods are general methods while the last two use the structure of the problem at hand, a system of first order linear differential equations, and the additional assumption of constant rates within the time-step. Intuition would be that tailored methods might be more efficient than using more general methods. However, solving a matrix exponential can also be time consuming, as time increases with the square of the dimension of the matrix. Also numerical problems abound,[78] Gallivan et al. published an algorithm specifically tuned to the case of Markov disease models.[79] Time demand can be further reduced by splitting up the matrix into separate matrices for independent clusters of diseases (see Technical Supplement A), where in practice many of these clusters consist of a single disease for which the matrix exponential can be solved exactly. DYNAMO-HIA therefore uses the third method, the algorithm of Gallivan et al.[79] for calculating the matrix exponential for clusters of more than one disease, and the exact solution for single diseases (both *simple* and with a cured state).

4

Introducing DYNAMO-HIA: A dynamic modeling tool for generic health impact assessments[†]

Abstract

Background Currently, no standard tool is publicly available that allows researchers or policy makers to quantify the impact of policies using epidemiological evidence within the causal framework of health impact assessment (HIA). A standard tool should comply with six criteria to be useful in the applied setting of HIA. With DYNAMO-HIA, we introduce such a generic software tool specifically designed to facilitate quantification in the assessment of the health impacts of policies.

Methods and Results

DYNAMO-HIA quantifies the impact of user-specified risk factor changes on multiple diseases and in turn on overall population health, clearly comparing one reference scenario with one or more intervention scenarios. The Markov-based modeling approach allows for explicit risk factor states and simulation of a real-life population. A built-in parameter estimation module ensures that only standard population-level epidemiological evidence, i.e. data on incidence, prevalence, relative risks, and mortality is required. DYNAMO-HIA provides a rich output of summary measures (e.g. life expectancy and disease-free life expectancy) and detailed data (e.g. prevalences and mortality/survival rates) by age, sex, and risk factor status over time.

[†]Stefan K. Lhachimi, Wilma J. Nusselder, Jet Smit, Pieter van Baal, Paolo Baili, Kathleen Bennett, Esteve Fernández, Margarete C. Kulik, Tim Lobstein, Joceline Pomerleau, Johan P. Mackenbach, Henrik C. Boshuizen (submitted for publication)

DYNAMO-HIA is controlled via a graphical user interface and is publicly available from the internet ensuring general accessibility. We illustrate the use of DYNAMO-HIA with two example applications: a policy causing an overall increase in alcohol consumption and quantifying the disease-burden of smoking.

Conclusion By combining modest data needs with general accessibility and user friendliness within the causal framework of HIA, DYNAMO-HIA is a potential standard tool for health impact assessment based on epidemiological evidence.

Introduction

Health impact assessment (HIA) is a combination of procedures, methods, and tools that judges the effect of a (planned) program, project, or policy on overall population health and the distributional effects within a population.[1] The rationale behind HIA is that many risk factors for chronic diseases are affected by policy measures outside the realm of health policy (e.g. transportation, food, or urban planning). Health impact assessments have been carried out at all governmental levels (e.g. local,[23] regional,[18] national,[24] and supranational[15]). The number of HIAs is likely to rise due to increased institutional adoption[25] and political will, in particular at EU level.[26] Currently, there is a diversity of approaches to the quantification of policy interventions.[2] However, for the quantification step in HIA a generic modeling tool – i.e. allowing for various and multiple chronic diseases and arbitrary risk factor – that takes into account the standard causal pathway assumed in HIA has been lacking.[62] The standard HIA causal pathway assumes that a policy intervention leads to a change in risk factor prevalence which in turn leads to changes in disease incidence and disease-related mortality.[9] The two objectives of HIA – to predict future consequences of implementing different options and to inform decision-makers in choosing between options[4] – address the technical core of quantification (predict) as well as the context (inform) in which an HIA takes place. Hence, a potential standard tool should aim for technical accuracy in the prediction of the effects of interventions on population health and yet be effective in the applied setting of an HIA, where time and resources are scarce. These objectives were operationalized into six criteria that a generic model should fulfill to be useful as a standard tool.[62] The first three criteria (real-life population, dynamic projection, and explicit

risk factor states) ensure that the model structure is sufficiently accurate in modeling changes in risk factor exposure over time in a real-life population in a transparent way. The last three criteria (modest data requirements, rich model output, and generally accessible) ensure a wide usability by accounting for the constraints of a decision-making process. This article proposes a software – DYNAMO-HIA (DYNAmic MOdeling for health impact assessment) – as a standard tool for the quantification of user-specified policy interventions within the HIA-framework.

Methods

Implementation of Requirements for a Standard Tool

We designed DYNAMO-HIA to satisfy the six criteria a generic standard tool for HIA should fulfill. DYNAMO-HIA models a closed real-life population, i.e. stratified by sex and age in 1-year age categories up to the age of 95 without migration (including the expected number of newborns). The model is dynamic in 1-year time-steps and projects reference and (several) intervention scenario(s) over time. DYNAMO-HIA has explicit risk factor states, i.e. at every time-step of the simulation each simulated individual is classified into a specific risk factor category. This ensures an accurate, unbiased estimation and increases the transparency of the simulation and the resulting output data.

DYNAMO-HIA has a parameter estimation module, mostly using methods taken from the RIVM-CDM, reducing data needs substantially.[66] Incidence and prevalence of a disease is only needed at the population level, i.e. specified by age and sex and not by each risk factor state. The module back-calculates the risk factor specific values using the relative risk from each risk factor state on diseases. The user can inspect these intermediate results when desired, thus improving transparency. DYNAMO-HIA provides rich simulation output available in three forms: first, raw output data allowing detailed analysis by age, sex, and risk factor status. This raw data either gives the cohort disease life table for every simulated cohort or the period data for every simulated year; second, several dynamic plots, e.g. population pyramids or survival rates, based on the data that contrast key information between the reference scenario and the intervention scenario; third, a range of summary outcome measures, e.g. cohort-, period-, or disease-free life expectancy. The

graphical user interface allows general accessibility; no programming or advanced computing skills are required.

Model Core

DYNAMO-HIA is a Markov-type model based on a multi-state model (MSM). The change of the state depends only on current characteristics (i.e. age, sex, risk factor status, and health status). The MSM is implemented as a partial micro-simulation combining a stochastic micro-simulation to project risk factor behavior with a deterministic macro approach for the disease life table.[80] In the micro-simulation module large numbers of distinctive risk-factor biographies are simulated: Given the age- and sex-specific transition probabilities between risk factor states, the risk factor status of each simulated individual is updated in annual increments (see Figure 4.1 for details). In the macro-module, as many disease life tables are constructed as there are risk-factor biographies. These disease life tables account for competing risks and multiple morbidity.[47] The exact configuration of the disease life tables, i.e. the number and kind of diseases, can be specified by the user (see Figure 4.2 for details). For every risk-factor biography, the probability of disease incidence and mortality over time is calculated, accounting for the current age, risk factor, and disease status (see Figure 4.3 for details). These biography-specific life tables are calculated for each birth-cohort, i.e. all individuals that are born in the same calendar year. For example, for a cohort of newborns, first, the risk-factor biographies are projected and then disease life tables are calculated. Older cohorts, i.e. born before the first simulation year, already start out having the disease prevalence as specified by the input data, which is then similarly updated. Population values are obtained by aggregating the individual biography/diseases life tables: either across cohorts at a given simulation time point to obtain period measures or along cohorts to obtain cohort specific measures (see Figure 4.4for details). The split into a micro- and a macro-module is done purely for computational convenience, micro- and macro-simulations yield the same result when used with the same data.[41,81] However, time and memory requirements in macro-simulations rise exponentially when the number of simulated states increases and micro-simulations – unlike customary multi-state life tables – do not require the a priori specification of all theoretically possible combinations of diseases/risk factor states, but only those states that are actually occupied. But for simulating

risk-factor biography	Age x	x+1	x+2	x+3	x+4
1	N	N	P	P	O
2	O	O	P	N	N
3	P	P	N	N	N
4	N	N	P	O	O
5	N	N	P	O	O
6	P	P	P	P	P
7	N	P	P	P	P
8	N	P	O	O	O
9	N	N	N	N	N
10	N	P	P	N	N

N= Normal Weight, P=Overweight, O= Obese

Figure 4.1: Example of risk-factor biographies for a risk-factor with three categories. DYNAMO-HIA simulates individuals and projects their risk-factor biographies. The risk-factor status is being updated in one-year increments, given age- and sex-specific transition probabilities. The age- and sex-specific risk-factor status determines the relative risk of a person to contract a disease or to die. DYNAMO-HIA allows one risk factor per scenario. This risk factor can be either categorical (up to ten categories), duration dependent (up to ten categories of which one is duration dependent, i.e. the risk on disease in this category depends on how long a person is in the category), or a continuous distribution (normal or log-normal, specified by entering mean, standard deviation, and, in the case of the log-normal, skewness).

rare events – e.g. lung cancer at young ages – micro-simulations require the simulation of large numbers of individuals, offsetting the savings in time and memory requirements.

The epidemiological model uses relative risks by risk factor class, i.e. incidences in exposed risk factor classes are a multiple of the incidence in the non-exposed. The total mortality, i.e. population level mortality by age and sex, is being decomposed in the mortality due to the diseases included in the model and other cause mortality. This decomposition assumes additive mortality: The total mortality rate in the population is explained as the sum of the mortality rate of the included diseases and other-cause mortality, i.e. mortality from all causes/diseases that are not explicitly included in the model.

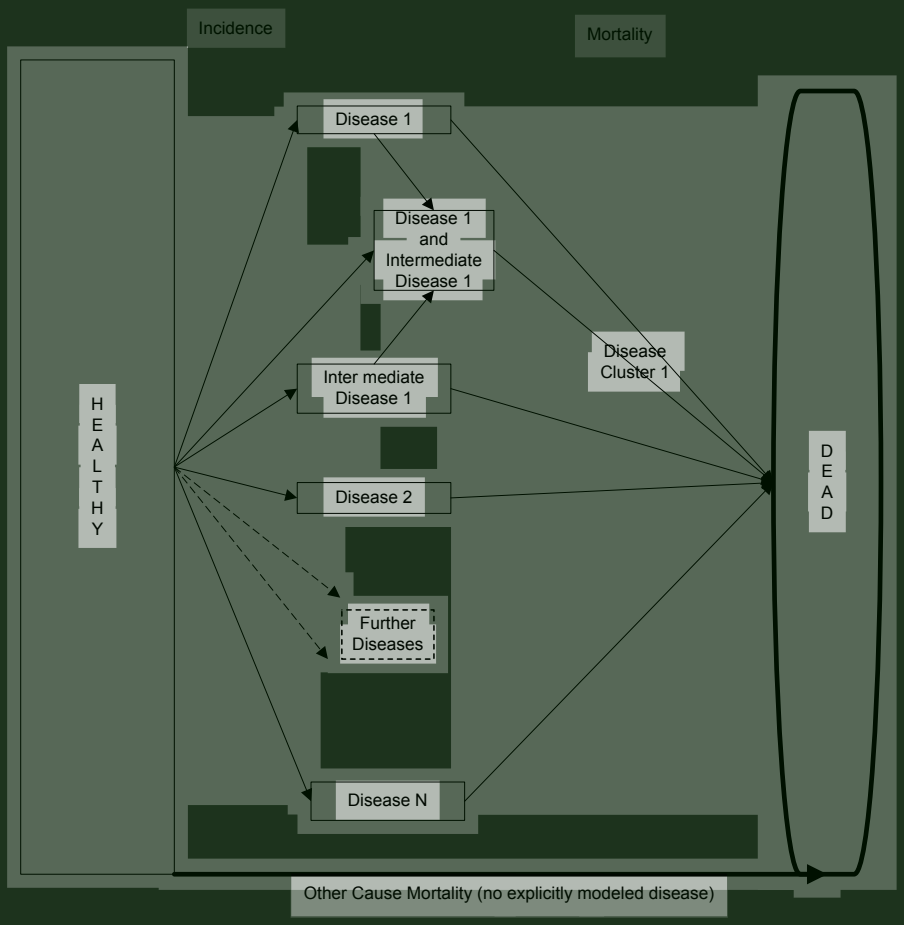

Figure 4.2: Stylized structure of disease life table. The disease life tables contain disease clusters. Each disease cluster consists of one or more diseases. Within disease clusters, intermediate diseases – that increase the risk of getting another disease – can be specified (e.g. having diabetes increases the risk of getting IHD). All diseases are chronic diseases, i.e. excess mortality depends on age and sex and not on time since onset of disease. However, acutely fatal and/or cured fraction can be specified for diseases. The disease life table assumes independence between disease clusters. The user can freely specify the relative risks from risk factor to disease, from risk factor to death, and from intermediate disease to other diseases.

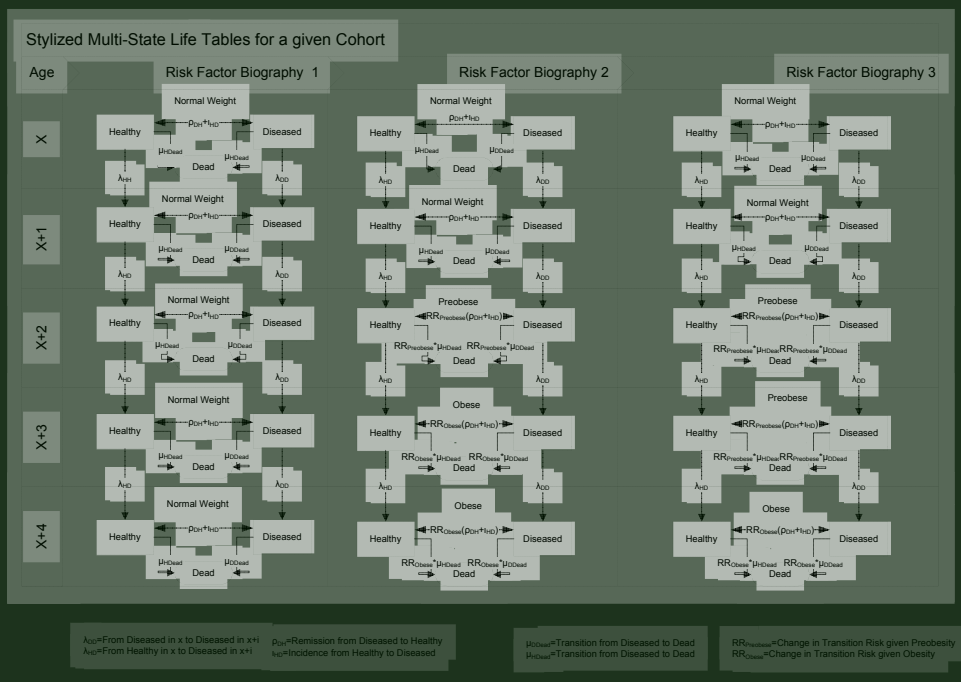

Figure 4.3: Stylized cohort life tables (with only one disease, three different biographies, and five time-steps). For every risk-factor biography, a disease life table is constructed. Disease incidence, i.e. transition from healthy to a disease, equals the baseline incidence, i.e. incidence when in a risk-factor class with a relative risk of one for the particular age- and sex-category, times the relative risk due to the given risk-factor and, in the case of an intermediate disease, diseases status. The transition from healthy to dead equals the baseline other-cause mortality of the healthy, i.e. age- and sex-specific total mortality rate minus the excess mortality rate of the diseases included in the disease life table, multiplied by the relative risk due to the given risk-factor status on other cause mortality. The transition from diseased to dead equals the sum of the excess mortality of the disease (given age and sex) and the baseline other cause mortality of the healthy, multiplied by the relative risk in the given risk-factor status. Remission is not explicitly modeled, but for diseases with cured fraction the excess mortality is zero in a "cured", i.e. user-specified, fraction. Partly acutely fatal diseases, i.e. diseases with very high mortality immediately after contracting the disease while for those who survive this critical period the excess mortality only depends on age and sex, are modeled by specifying the fraction of the incidence cases that die immediately.

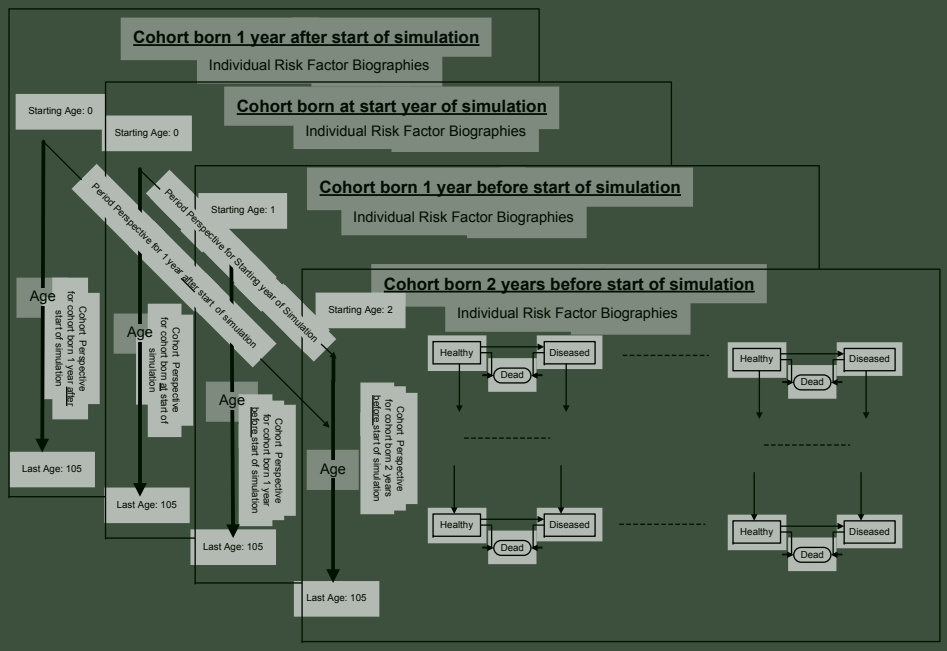

Figure 4.4: Schematic overview of the dimension of a multi-cohort, multi-state life table. Each plane is a distinct cohort with varying starting ages for cohorts already existing at the starting year of the simulation and starting age zero for cohorts born during the simulation run. The cohort life tables, consisting of the set of individual risk-factor biographies, follow every already existing birth cohort until the cohort reaches 105 years of age. In addition, every year of the simulation a cohort of newborns is created and – after simulating individual risk-factor biographies for them – is followed through the appropriate disease life tables as well. This allows collecting health data for each cohort according to their risk-factor status (longitudinal) or the health status of the population by age, sex, and risk-factor status by each year of the simulation (cross-sectional).

Modeling policies with DYNAMO-HIA

The goal of HIA is to compare the effect of several policies/interventions on future population health, keeping the status quo as the reference scenario. Within DYNAMO-HIA, policies can be modeled in two ways (both approaches can be applied simultaneously and/or targeted at selected parts of the population only). The first approach is to define a counterfactual risk factor prevalence that is assumed to be reached after a successful one-time, sustained intervention, e.g. a reduction in alcohol consumption caused by a tax increase or a ban on consumption in public. The approach of defining counterfactual risk factor prevalences is akin to epidemiological methods, where total or partial eradication of a risk factor is quantified. DYNAMO-HIA does this quantification dynamically, i.e. effects are projected over time. The second approach is to alter the transition probabilities between different risk factor states, i.e. changing the risk factor behavior of the population. This approach is closer to the reality of many health interventions that try to influence life style choices of individuals, e.g. halving the future number of teenagers that become obese. The specification of the transition probabilities influences greatly the future development of the risk factor prevalence, which is always debatable. As an option DYNAMO-HIA provides the use of net-transition probabilities: DYNAMO-HIA estimates internally the transition probabilities that keep the age-specific risk factor prevalence constant, ignoring any future cohort effects.

Illustration

To illustrate the usability of DYNAMO-HIA, we present two stylized example applications. The first illustration projects the consequences of a policy-induced increase in alcohol consumption and resembles a prospective HIA. The second illustration quantifies the changes in population health when smoking would be eradicated and resembles a burden of disease study. In both applications, we model the effect of risk factors on total mortality and nine diseases – ischemic heart disease (IHD), stroke, diabetes, chronic obstructive pulmonary disease (COPD), breast-, lung-, esophageal-, colorectal-, and oral-cancer – and keep the age-specific risk factor prevalence constant over time by using net-transition probabilities between risk factor classes, i.e. ignoring any future cohort effects. Hence, the difference between the reference scenarios and the intervention scenarios depends solely on the different

initial risk factor prevalences. The data sources and the relative risk used are shown in detail in the Appendix.

Liberalizing access to alcohol: The Swedish case

In 2004, Sweden had to lift her ban on private alcohol imports.[82] Prospective studies were forecasting an increase in overall alcohol consumption and, consequently, a worsening of a number of alcohol-related harm indicators. In our reference scenario, we keep the alcohol consumption prevalence observed in 2002 constant during the projection period and assume a one-time change in the consumption of pure alcohol by 1L per capita, producing a counterfactual risk factor prevalence as seen in Figure 4.5 for the intervention scenario. We project both scenarios for 25 years in the future (see Table 4.1). The annual excess number of deaths due to increased alcohol consumption is on average approx. 170 deaths, accruing to some 4,300 additional deaths over the 25-year-period. This projected difference in overall population mortality also reflects all other effects a risk factor has on other-cause mortality accounting for not included diseases and – more salient in the case of alcohol – injuries/accidents via the relative risk of a risk factor on total mortality. This absolute number is rather small compared to the overall population of some 9 million; hence, the effect on total life expectancy and, similarly, the overall difference in disease-free life expectancies between the reference and the intervention scenario are negligible.

Alcohol intake has a pronounced effect on a number of diseases that are projected by the model. In projection year 25, the biggest difference in absolute cases is for diabetes with approx. 6,600 more cases, followed by stroke with an excess prevalence of approx. 1,700 cases. Ischemic heart disease, the most prevalent of the included diseases, is overall less affected by the change in alcohol intake. The population prevalence differs only marginally over the simulation period, but still accounts for approx. 700 less cases; this is partly caused by the beneficial effect of moderate drinking for some age groups (see Table C.8 for the corresponding relative risks). From the five included cancers, the increase in breast cancer is the most notable: in projection year 25, the excess prevalence is approx. 1,800 cases in the intervention scenario. For the other cancers, the increases in prevalence cases are relatively minor: for oral cancer approx. 750, for colorectal cancer approx. 280 cases, and for esophageal cancer approx. 60 additional cases in the counterfactual scenario. The

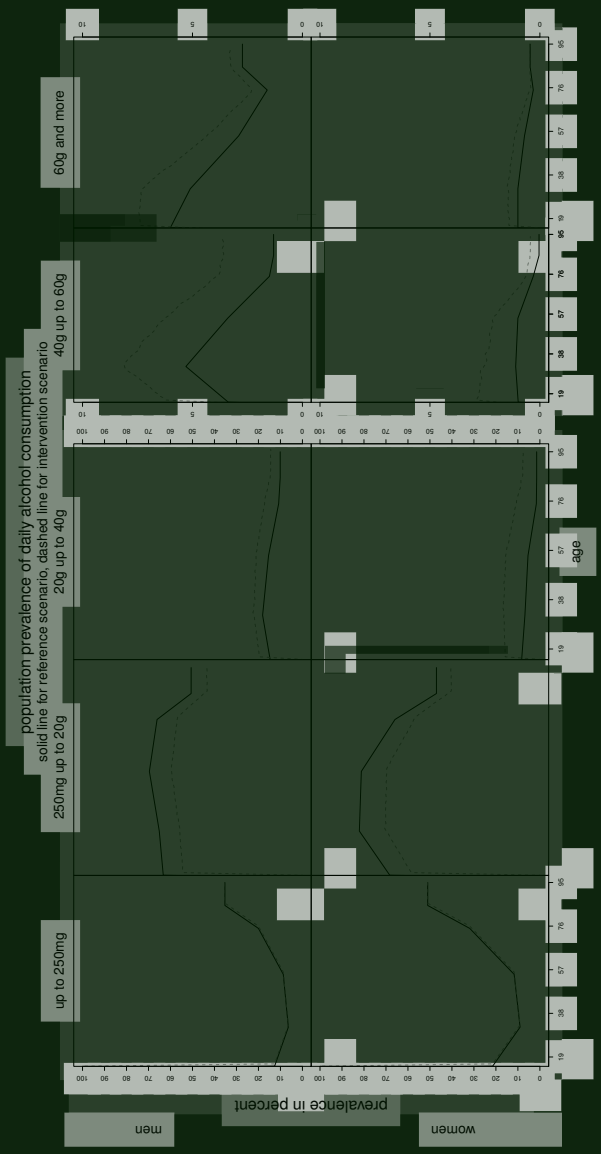

Figure 4.5: Swedish prevalence of alcohol consumption intervention scenario compared with reference scenario (Alcohol consumption is measured by five categories of daily intake of grams of pure alcohol: 0 - <0.25g/d, 0.25 - <20g/d, 20 - <40g/d, 40- <60g/d, ≤60g/d).

number of COPD cases show a slight decrease although it is not causally related to alcohol intake. This is due to the higher number of deaths, thus there are less persons alive to contract this disease.

Total elimination of smoking: A projection with UK data

Smoking is a major public health concern. This illustration quantifies the gain in population health obtainable if an entire population would consist of never smokers compared to a real-life population that keeps the currently observed smoking behavior unchanged. Smoking is measured in three categories (never-, former-, current-smoker). The data for this illustration are from the UK and projected 25 years into the future (see Table 4.1 and 4.2). In the counterfactual, the whole population consists of never smokers and no uptake of smoking. After 25 years, the population of never-smokers is projected to have approx. 1,510,000 more individuals than a population keeping the current smoking behavior. This translates into a total life expectancy of 81.4 years for the counterfactual compared to 79.0 years for the reference scenario. This gain in life expectancy is substantially larger for men than for women: For men the difference is more than 3 life years (76.9 in the reference scenario compared to 80.0 in the intervention scenario) and women 1.8 years (81.0 compared to 82.8). The projected life expectancies clearly demonstrate that in DYNAMO-HIA no autonomous trends are assumed, e.g. a secular increase in life expectancy that one may expect over the next 25 years. Smoking also has a causal effect on a number of diseases. The biggest reduction in the modeled diseases is for COPD. In projection year 25, the average life years lived with COPD is approx. 0.9 years less in the intervention scenario than in the reference scenario, more than halving population prevalence from 1.7% to .5%. The next biggest reduction is for IHD with approx. half a year less expected life years with this disease, a difference in prevalence of 1 percentage point. Similarly, the prevalence of stroke goes down by approx. 0.4 percentage points (from 2.3% to 1.9%). The three included cancers that are related to smoking are reduced as well (lung cancer by approx. 87,000 cases, esophageal cancer by approx. 38,000, and oral cancer by approx. 9,200 cases, respectively). However, other included diseases that are not causally related to smoking (diabetes, breast-, and colorectal cancer) increase in prevalence thanks to the larger number of surviving individuals that are now at risk of contracting those diseases.

Table 4.1: Number of disease cases and population prevalence (in percent) for example applications.

	Swedish Alcohol Example					UK Smoking Example				
		Reference Scenario		Intervention Scenario			Reference Scenario		Intervention Scenario	
	year	numbers	percent	numbers	percent	year	numbers	percent	numbers	percent
IHD	1	354,747	3.9	354,747	3.9	1	2,183,447	3.6	2,183,447	3.6
	25	428,727	4.7	428,026	4.6	25	2,666,917	4.4	2,144,465	3.5
Stroke	1	150,271	1.7	150,271	1.7	1	1,002,594	1.7	1,002,594	1.7
	25	192,924	2.1	194,616	2.1	25	1,374,698	2.3	1,156,176	1.9
Diabetes	1	368,787	4.1	368,787	4.1	1	1,559,679	2.6	1,559,679	2.6
	25	385,216	4.2	391,793	4.3	25	1,850,392	3.1	1,999,847	3.2
Lung Cancer	1	4,613	0.1	4,613	0.1	1	82,082	0.1	82,082	0.1
	25	5,753	0.1	5,750	0.1	25	107,393	0.2	20,334	0
Oral Cancer	1	9,377	0.1	9,377	0.1	1	55,804	0.1	55,804	0.1
	25	11,738	0.1	12,495	0.1	25	70,949	0.1	33,013	0.1
Esophageal Cancer	1	971	0	971	0	1	11,231	0	11,231	0
	25	1,241	0	1,300	0	25	14,535	0	5,324	0
Colorectal Cancer	1	36,415	0.4	36,415	0.4	1	248,380	0.4	248,380	0.4
	25	47,775	0.5	48,062	0.5	25	329,116	0.5	370,596	0.6
Breast Cancer	1	90,444	1	90,444	1	1	543,738	0.9	543,738	0.9
	25	108,854	1.2	110,661	1.2	25	670,013	1.1	705,919	1.1
COPD	1	105,052	1.2	105,052	1.2	1	525,247	0.9	525,247	0.9
	25	131,118	1.4	130,850	1.4	25	1,016,422	1.7	278,194	0.4
With at least one disease	1	918,921	10.2	918,921	10.2	1	5,148,112	8.6	5,148,112	8.6
	25	1,081,720	11.7	1,088,547	11.8	25	6,726,107	11.1	5,792,303	9.4
Size of total population	1	9,002,148		9,002,148		1	59,987,010		59,987,010	
	25	9,210,437		9,206,131		25	60,416,567		61,929,848	

Table 4.2: Period based total life expectancy and expected number of years with a disease for the UK example application

		Women			Men		
	year	Reference Scenario	Difference between Scenarios	Inter-vention Scenario	Reference Scenario	Difference between Scenarios	Inter-vention Scenario
total life expectancy	1	81.3	0.2	81.5	77	0.8	77.8
	25	81	1.8	82.8	76.9	3.1	80
IHD	1	3.3	0	3.3	4.3	0.2	4.5
	25	3	-0.4	2.6	3.8	-0.6	3.2
Stroke	1	1.9	0	1.9	1.8	0.1	1.9
	25	1.8	-0.2	1.6	1.8	-0.2	1.6
Diabetes	1	2.3	0	2.3	2.6	0.1	2.7
	25	2.1	0.1	2.2	2.5	0.3	2.8
Lung Cancer	1	0.1	0	0.1	0.2	0	0.2
	25	0.1	-0.1	0	0.2	-0.2	0
Oral Cancer	1	0.1	0	0.1	0.1	0	0.1
	25	0.1	-0.1	0	0.1	-0.1	0
Esophageal Cancer	1	0	0	0	0	0	0
	25	0	0	0	0	0	0
Colorectal Cancer	1	0.4	0	0.4	0.5	0	0.5
	25	0.4	0.1	0.5	0.5	0.1	0.6
Breast Cancer	1	1.7	0	1.7	n/a	n/a	n/a
	25	1.7	0.1	1.8	n/a	n/a	n/a
COPD	1	0.8	0	0.8	0.9	0.1	1.0
	25	1.3	-0.9	0.4	1.2	-0.9	0.3
With at least one disease	1	8.9	-0.8	8.8	8.4	0.3	8.7
	25	8.7	-0.8	7.8	8.2	-0.9	7.3

Discussion

Within the rapidly developing field of HIA, no standard method on quantification has emerged yet,[11] but three approaches predominate the field: regression-based methods, quantitative risk assessment, and population health models. The regression-based methods originate in econometrics and usually estimate the long term relationship between exposure (e.g. per capita consumption) or proxy variables (e.g. tax rate on alcohol) and health outcomes of interest on an aggregate level adjusting for further variables as suggested by (economic) theory. This approach usually takes only limited notice of underlying epidemiological mechanisms. Quantitative risk assessment originating from (environmental) exposure assessment of toxic substances makes explicit use of dose-response relationships derived through epidemiological studies. These approaches are usually static, i.e. not accounting for changes over time in real-life populations. Population health models combine epidemiological evidence and insights on causality to dynamically quantify the effect of risk factors on population health.

DYNAMO-HIA fills a gap among the already existing population health models that are suggested for application in HIA.[62,83] Compared to existing models, DYNAMO-HIA strikes a balance between being sufficiently technically accurate and ensuring wide usability. Technically equal or more complex models – e.g. POHEM, ARMADA, RIVM-CDM – allow a greater flexibility in modeling, but are not publicly available and they require highly specialized input data, and, moreover, proficiency in specialized programming languages (except ARMADA). More accessible models – e.g. PREVENT, Proportional Multi-state Life Table (MSLT), GBD – lack dynamic projection capabilities (except PREVENT and multiple cohort versions of the MSLT[84]) and do not have explicit risk factor states, this technical simplification ignores mortality selection and may lead to biased estimates.[62]

DYNAMO-HIA is specially designed to fit within the standard framework of HIA synthesizing elements of already well-established modeling approaches. Our approach allows for a flexible risk factor configuration (categorical, duration dependent, continuous), generic chronic diseases as specified by the user (with intermediate diseases, partially fatal diseases, and/or diseases with a cured fraction), arbitrary specification of age- and sex-specific – relative risks, and minimal data needs by requiring only population level data. Furthermore, a mouse-driven graphic user interface allows straightfor-

ward handling of the software, i.e. no knowledge of a programming language is required. In addition to exporting the existing, partly customizable, graphs into files – e.g. detailed plots of mortality rates or prevalences of risk factors or diseases, both over time and age-specific – most calculated data can be exported for use in separate software (e.g. Excel). This raw output data allows further analysis such as grouping diseases into categories (e.g. IHD and stroke or all cancers), to include costs, or to construct graphs.

DYNAMO-HIA simulates the effect of a single risk factor on a population without migration. However, the categorical risk factor can be used to partition the population in up to ten distinctive categories. For example, a population could be partitioned along a risk factor – say, never-smokers and smokers – and socioeconomic status – say, with and without college education – having in total four different groups assessing a policy that is more successful for people with certain socioeconomic status. The possibility of partitioning a population also allows to quantify the effect of an environmental hazard. In this case, for example, the population is partitioned according to their proximity to the hazard source – say, noise exposure or air pollution due to a new airport – with 5% of the total population living less than 5km from the hazard source, 5% to 10% living less than 10km and so on. This requires, of course, sufficient insight into which part of the population is affected and knowledge of the relative risks of the modeled exposure on the included diseases and total mortality.

A category may also stand for a combination of known risk factors: For example, smoking status (smoking/non-smoking) and BMI (normal-/overweight/obese) could be modeled by partitioning the population into six distinctive risk factor categories. However, it requires knowledge about the relative risk of the combined risk factor class – say, relative risk of being obese and a smoker on the included diseases and total mortality.

The overall performance of a model crucially depends on the quality of the input data. In particular for dynamic models, the epidemiological data has to be mutually consistent,[69,85] otherwise projected changes in the prevalences might be caused by mismatching data and not by the changes in the risk factors. A limitation is that an autonomous trend in the rates, e.g. annual reduction in overall mortality or disease incidence, cannot be specified. Autonomous trends are often observed for past time periods and caused by a number of factors; chief among them are improved curative interventions and changed risk factor behavior. In a risk-factor based model, however, the

specification of a future autonomous trend must be net of any underlying risk factor behavior as this is already specified explicitly at some other place in the model. Such specific data on future trends is hardly reliably available, if at all, and would in most cases only modestly affect the difference between reference and intervention scenarios. Hence, an ordinal ranking of policy alternatives would be rarely affected and still reveal the most effective intervention.

In health impact assessment, three criteria are used to assess validity: formal validity, plausibility, and predictive validity.[86] *Formal validity* assesses the degree to which correct methods are applied correctly. The model structure of DYNAMO-HIA is well founded in epidemiological evidence – incidence, prevalence, and excess mortality – and demographic modeling practice, i.e. a multi-state Markov-type model of chronic disease with explicit risk factor states and inclusion of intermediate diseases. Plausibility assesses the degree to which an observer deems that the theoretical framework is understandable, applicable, and plausible. Hence, DYNAMO-HIA deliberately restricts itself on the well-established causal chain *risk factor exposure* $\rightarrow$ *incidence* $\rightarrow$ *prevalence* $\rightarrow$ *disease-related mortality* $\rightarrow$ *overall population health* and requires only data that is available in sufficient quality for the most common diseases (e.g. cancer, CVD, diabetes, COPD) and risk factors (e.g. smoking, BMI, alcohol) in developed countries. In the Swedish example application, our results for the number of excess death is slightly lower than estimates based on a regression approach utilizing historical relationships and aggregate-data pooled from several Nordic countries.[82] One reason for this difference lies in the relative risks on all-cause mortality used in our illustration. Those are taken from epidemiological studies and capture only the effect of individual exposure, i.e. drinking behavior. Consequently, our results do not account for broader effects that a change in alcohol consumption has on population health, i.e. abstainers or moderate drinkers become victims of increasing alcohol-induced violence or accidents, caused by the increased number of intoxicated drinkers.

Plausibility and well-established formal methods should not be mistaken with constantly delivering expected results. Dynamic projections may reveal counterintuitive, yet plausible results and, hence, lead to important insights. In the smoking application, for example, the number of breast cancer cases in the never-smoker scenario is larger than in the reference scenario although smoking has – in this application – no causal link to breast can-

cer incidence. This seemingly unexpected result is caused by an increase in overall longevity of a healthier living population and, hence, increased number of females that are susceptible to breast cancer. This phenomenon is well known among modelers of health care costs, where dynamic analysis showed repeatedly that a population level reduction in obesity or smoking may lead to higher health care costs in the long run.[52,87]

Predictive validity is the degree to which predictions are confirmed by facts; Veerman et al, however, list several reasons why this criterion usually cannot be established in the context of HIA. We emphasize that a software model like DYNAMO-HIA is always a decision-support tool only. It helps to quantify the expected differences in population health given two (or more) different scenarios: one of them a baseline scenario (without the intervention) and one (or more) scenario(s) with intervention(s). It does not predict the development of future population health as such. Decision-makers must be constantly aware that real-world phenomena are necessarily more complex and no model can predict future events accurately. In HIA it might be useful to avoid calling the results of mathematical models "predictions", but rather projections of *what-if* scenarios in a clearly defined and simplifying framework. The term "prediction" should be reserved for the entire process, in which a software model is only one element of the used evidence.[62,88,89]

Internal validity was extensively tested. To allow future thorough checking of cross validity also by outside experts, the software and the source code are publicly available (`www.DYNAMO-HIA.eu`). DYNAMO-HIA allows in its current form for unproblematic one- and multi-way sensitivity analysis by easy manipulation of all input parameters. Like most other population health models, however, the current version of DYNAMO-HIA does not include a probabilistic sensitivity analysis (PSA). Implementing a PSA in population health models is time and cost intensive. Moreover, the extra data needed to conduct a PSA is difficult to obtain and preparing them requires expert knowledge. However, DYNAMO-HIA can be used in batch mode, allowing users with sufficient computing skills to build a PSA shell around the software when desired.

DYNAMO-HIA is available for free download. Furthermore, DYNAMO-HIA includes a data set covering a large number of EU-countries. This internally consistent data set has prevalence data for three risk factors (smoking, BMI, alcohol), nine diseases (incidence, prevalence, excess mortality), and population data (e.g. total mortality, projected number of newborns). This

data set allows instant use of DYNAMO-HIA for the covered countries. However, DYNAMO-HIA is also usable with external data on other countries, (sub-)populations, disease, or risk factors. Furthermore, the already included data set can be easily updated when more recent data becomes available.

Conclusion

DYNAMO-HIA differs from other population health models for HIA[83] in several important aspects. From the outset, it has been designed for public use within HIA-applications by featuring a user-friendly graphical interface and employing a model structure that ensures accurate simulation using epidemiological evidence while having modest data needs.

Part II
Selected Applications

6

Health impacts of increasing the European Union-wide excise duty on alcohol: A dynamic projection[†]

Abstract

Background Western Europe has high levels of alcohol consumption, with corresponding adverse health effects. There is clear evidence that increased alcohol prices reduce alcohol consumption. Currently, a major revision of the EU excise tax regime is under discussion. We seek to quantify some of the potential health consequences of increasing price on alcohol by varying tax rates.

Data and Methods Using alcohol consumption data for eleven member states covering 80% of the EU-27 population, and corresponding country-specific disease data (incidence, prevalence, and case-fatality rate of alcohol related disease), we projected the expected changes in selected measures of population health that might arise from changes in alcohol price.

Results Even a modest price increase of 20% leads to fewer cases of stroke, diabetes and cancer, and fewer deaths in both men and women. Effects are larger in men. An increase in alcohol prices towards those currently in Finland (the highest in the EU) would postpone approx. 54,000 male and approx. 26,000 female deaths. Moreover, the prevalence of a number of chronic diseases would be reduced, in men by approx. 97,850 individuals with stroke,

[†]Stefan K. Lhachimi, Katie Cole, Jet Smit, Pieter van Baal, Paolo Baili, Kathleen Bennett, Esteve Fernández, Margarete C. Kulik, Tim Lobstein, Joceline Pomerleau, Martin McKee, Wilma J. Nusselder, Johan P. Mackenbach, Henrik C. Boshuizen (submitted for publication)

65,850 with diabetes, and 62,000 with cancer, and in women by approx. 19,100 (stoke), 23,450 (diabetes), and 27,000 (cancer), respectively.

Conclusion Curbing excessive drinking throughout the EU completely would lead to substantial gains in population health, with harmonization of prices to the level seen in Finland achieving some of those gains. The dynamic modeling tool DYNAMO-HIA, despite its inherently conservative estimates, can inform the current debate on revision of tax rates in the EU and supports a substantial price increase.

Introduction

The European Union has an alcohol problem.[90] The 2004 Global Burden of Disease Study estimated that, in EURO-A (corresponding to Western Europe), 11.1% of the burden of disease among men was attributable to alcohol, as well as 1.6% of the burden among women, while the corresponding figures for EURO-B (central Europe) were 10.2% among men and 2.5% among women.[100] Although alcohol-related deaths have been falling in some countries, such as France, they have been rising rapidly in others, such as the United Kingdom.[101]

At a population level, the major determinants of alcohol consumption are access (in terms of density of outlets, opening hours, and age restrictions on purchases), advertising, and price. There is now an extensive volume of research on the last of these, with the most recent meta-analysis reporting that a 10% increase in price is associated, on average, with a 4.6% reduction in consumption of beer, with the corresponding figures 6.9% for wine, and 8% for spirits.[102] Historically, in most European Union countries taxation of alcohol has been used primarily as a means of raising revenue, with the health effects somewhat as an afterthought (the Nordic countries are exceptions[103]). However, this is now changing, exemplified by the recent unsuccessful attempt by the Scottish government to introduce a minimum price per unit of alcohol, drawing on research undertaken in England showing that this would be highly effective,[104] especially in the face of massive discounting by supermarkets, many of which sell alcohol at below cost price as loss-leaders.

Within the EU, overall taxation levels are determined nationally. There is a binding minimum excise duty rate but this was last agreed in 1992 and is a fixed monetary sum per volume of pure alcohol. Its value has long since been eroded by inflation. Simply to adjust for inflation, this rate would have to rise

Table 6.1: Data sources for alcohol consumption data.

Country	Name of survey, Year	Number of respondents	Age range of raw data*
Denmark	Danish National Health Interview Survey, 2005	N=14,468	16+
Finland	Finnish Drinking Habits Survey, 2008	N=2,725	15-69
France	Enquete Nationale Nutrition Sante (National Health and Nutrition Survey), 2006-2007	N=2,640	20-74
Germany	German Epidemiological Survey of Substance Abuse (ESA), 2006	N= 7,571	18-64
Ireland	Survey of Lifestyles, Attitudes and Nutrition in Ireland (SLAN), 2007	N=7,964	18+
Italy	Everyday Life in 2007: Multipurpose Survey on Households	N=41,491	15+
Netherlands	Permanent Survey on Living Conditions (POLS) – Health Interview Survey, 2005/2006/2007	N=17,93	15+
Poland	National Multicenter Health Survey (WOBASZ), 2005	N=13,256	20-74
Spain	Spanish National Health Survey (SNHS), 2006	N=28,628	16+
Sweden	The (Alcohol) Monitoring Study, 2002	N=˜18,000	16-80
UK	UK General Household Survey, 2006	N=13,503	16+

*Higher ages have been extrapolated using SHARE data; see Method section for details.

Table 6.2: Values for price level indicator (PLI) in 2009 and factor to reach highest price level.

Country	PLI	Price increase necessary to reach Finnish price level
EU-27 countries	100	70.0%
DK	134.8	26.0%
DE	90.5	88.0%
IE	166.9	2.0%*
ES	84.3	101.0%
FR	95	79.0%
IT	112.4	51.0%
NL	98.9	72.0%
PL	89	91.0%
FI	169.8	n/a
SE	137.6	23.0%
UK	117.2	45.0%

Source: Eurostat *set to zero in the respective scenario

by 44%. Decisions about national tax levels are, however, influenced by concerns, often encouraged by the alcohol industry, about the ability of individuals to transport large quantities of alcohol freely across borders for personal use, as when the Finnish government reduced taxes in response to pressure from the domestic alcohol industry following the accession of neighboring Estonia, where prices are much lower, to the European Union. Nevertheless, prices vary markedly, with the highest prices in Finland, at 70% above the EU average, and the lowest in Romania, at 30% below it.

The taxation regime for alcohol within the EU is at last being revisited. However, the discussions are being driven in large part by considerations of the impact of different tax regimes on the functioning of the internal market rather than on public health. We argue that the health consequences of any changes should at least be considered but this is not easy in the absence of information on what might happen as a result of any changes. Previous attempts to quantify this have largely been restricted to one country.[105–107] For this reason, we describe the results obtained from a newly developed dynamic modeling tool that can estimate the effect of a range of alcohol price increases on selected chronic diseases and on total mortality in the EU population. We have been able to obtain data from eleven countries, covering some 80% of the EU population, and can look at the effect of different price regimes for nine chronic diseases: IHD, stroke, diabetes, COPD, breast-, lung-, esophageal-, oral-, and colorectal-cancer.

Methods

The analysis was undertaken using the dynamic population health modeling tool DYNAMO-HIA which combines three pieces of evidence: a) country-specific data on patterns of alcohol consumption in the EU to determine population-level exposure; b) price elasticity of alcohol demand to predict the expected change in alcohol consumption when prices increase; and c) country-specific data on current population structure and projected births, disease-incidence, -prevalence, and -mortality, the corresponding relative risks, and on total mortality.

Alcohol consumption prevalence

In developing the DYNAMO-HIA data base, we obtained age- and sex-specific prevalence data on alcohol consumption for eleven EU-member states, covering 80% of the EU-27 population. These data is derived from population-based surveys of individual self-reported consumption (Table 6.1). The original data were collected in various formats, although most often either quantity-frequency or within dietary surveys using food frequency or period recall methods. These were used to derive consumption in grams of pure alcohol consumed per person per day (g/d), using national conversion factors where available but otherwise estimating a "standard" drink as 12g of ethanol. We attempted to adjust for under-reporting by combining survey and sales data but identified problems with this approach that, as far as we know, have not previously been reported,[108] so that we generated implausible results.[109]

For both sexes, the same five consumption levels were categorized: <.25g/d, .25-20g/d, 20-40g/d, 40-60g/d, and over 60g/d. Following the standard approach taken in earlier modeling exercises, a uniform distribution of consumption was assumed within each category, giving a category-specific mean consumption of .125g/d, 9.875g/d, 30g/d, 50g/d, and 90g/d, respectively.[108] It was necessary to transform the available data into one-year-age intervals. To do this, the available data were smoothed and, where data were missing for higher age groups, they were extrapolated using data on the rate of change in consumption with age derived from the Survey of Health, Ageing and Retirement in Europe (SHARE). For details see Table 6.1, Figure 6.1, and the corresponding work package report.[109]

Price elasticities

The effect of price changes on consumption is modeled via price elasticities. Price elasticity measures the average proportional reduction in consumption when the price of a commodity increases. Those elasticities are estimated empirically. Estimates of price elasticity of alcohol consumption vary and in individual studies seem to be influenced by a range of variables – such as the country studied, the type of alcohol (beer, wine, or spirits), and, at an individual level, age and current consumption level. However, it was not possible to produce a generalizable formula that could be used to calculate specific elasticities in the diverse settings we were studying. Hence, as a recent

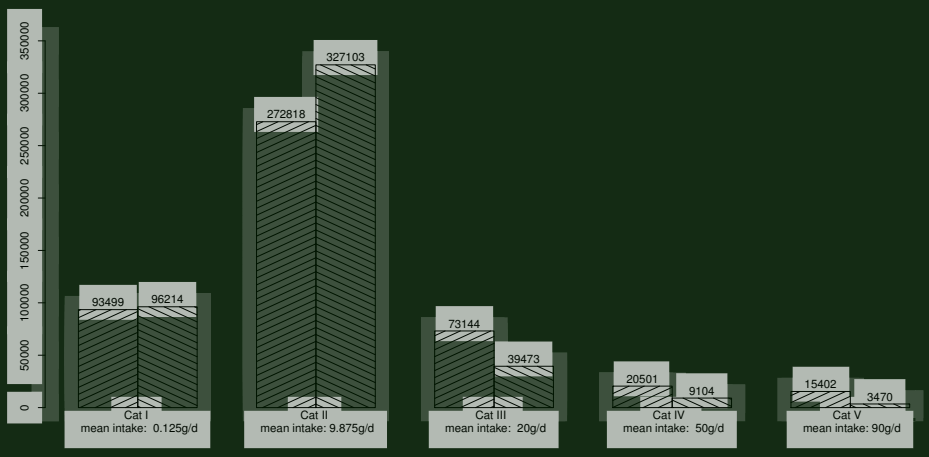

Figure 6.1: Example of change in age- and sex-specific alcohol consumption prevalence caused by a 88% price increase (absolute numbers are for 30 year old German males, left bar reference scenario, right bar "Finnish" scenario). The price increase leads to a decrease in Cat III to Cat V and an increase in Cat I and Cat II.

meta-analysis reported a fairly similar overall median elasticity for alcohol applicable across populations and beverage types of approx. -0.5, whereby a price increase of 20% leads to an average reduction in alcohol consumption of 10%, we used this figure.[110]

As noted above, the magnitude and precise structure of any future European excise regime are still under discussion. Given the different pricing policies available, levels of consumption, purchasing parities, and duty levels in EU-countries, it would be extremely complex to assess the health impact of absolute changes in price between different countries. Hence, we opted for a series of illustrative scenarios in which we applied a range of percentage increases in current alcohol prices. We used four different scenarios, comprising 20%, 40%, 60%, and 80% increases in alcohol prices. In addition, we estimated the effects of a "Finnish" scenario where we quantify for each country a new level of consumption that corresponds to adoption of the highest price level currently observed in the EU, that is Finland's price level (Table 6.2). In this scenario, every individual would have to spend the same amount to obtain an identical basket of alcoholic drinks.[111]

For each country in our analysis, we estimated five new alcohol consumption prevalences as a consequence of these price changes. Using the age- and

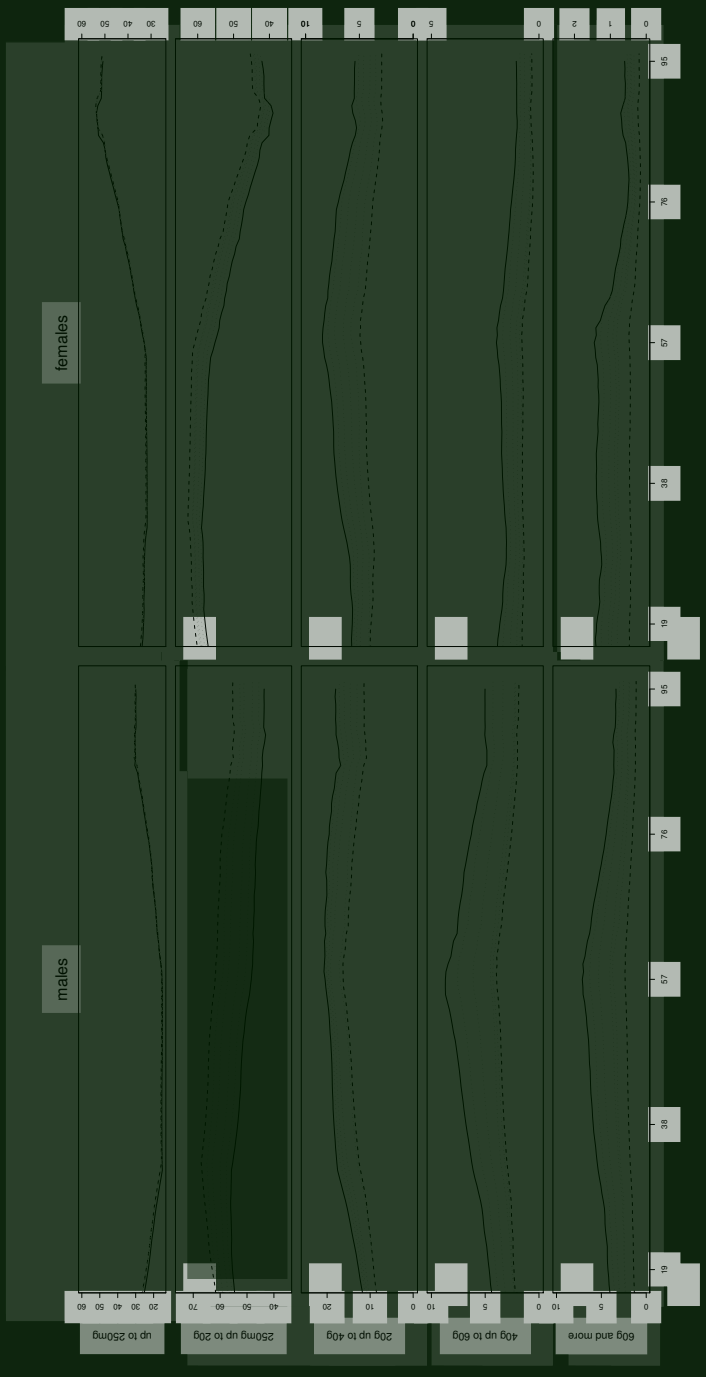

Figure 6.2: Alcohol population prevalence in percent for all countries combined (reference scenario black line, 80% scenario dashed line, light grey dotted lines are all other scenarios, i.e. 20% scenario, 40% scenario, 60% scenario, and Finnish scenario); y-scales vary by category.

sex-specific distribution across and within the five alcohol categories, we assigned every individual a corresponding daily alcohol consumption level and adjusted this according to the expected consequences of the price change, e.g. a 20% price increase with a price elasticity of -0.5 leads to a 10% reduction in individual consumption. Then we grouped the individuals back into the age- and sex-specific age categories according to their new consumption level (compare Figure 6.1). Where there were small changes in price, most individuals stayed in their current category. For example, if consumption was reduced from 33g/d to 30g/d, the category for this individual is in both cases 20-40g/d. Abstainers are only very marginally affected by the existing price level, as price is not likely to affect the share of drinkers significantly but rather the amount drinkers consume. Only approx. 1% of abstainers cite price as the main reason for abstaining. [112]

Finally, we constructed for each country a scenario whereby the age- and sex-specific share of abstainers stays constant but all current consumers of alcohol are light drinkers, i.e. a consumption of 250mg-20g/d. This scenario serves as the theoretical upper bound of the health improvements that a population-wide alcohol control policy can achieve through tax changes, i.e. only altering prices and not drinking ages or modes of drinking.

Epidemiological data

In developing the DYNAMO-HIA package, we compiled data on population structure and patterns of disease (see Appendix C for details). For each of the countries in our analysis, we use age- and sex-specific data on the population, i.e. size, projected birth numbers, and total mortality rate. Furthermore, we used country-specific data on nine major chronic diseases: IHD, diabetes, COPD, stroke, and lung-, breast-, colorectal-, oral-, and esophageal-cancer. Each disease is characterized in terms of age- and sex-specific incidence, prevalence, and excess mortality. For stroke and IHD, excess mortality is a combination of two factors: a) an age- and sex-dependent increase in individual mortality when having those diseases, and b) an acute but temporary increased mortality when the disease occurs, to reflect how mortality from these diseases is higher at their onset. Where appropriate, missing data were back-calculated using the DisMod II software. This utilized the mathematical relationship between incidence, prevalence, and excess mortality for chronic diseases within a given population. [69] In addition, DisMod II was used to en-

sure smooth and internally consistent data. Lastly, we use data on age- and sex-specific relative risks causally connecting alcohol consumption to disease incidence and all-cause mortality. Additionally, the increased relative risk of IHD- and stroke-incidence when having diabetes is included.

Dynamic Modeling

The DYNAMO-HIA package quantifies the effects of changing risk factor exposure on population health (see Chapters 3 and 4 for details). Specifically, it quantifies the effects of changes in risk factor prevalence on different health outcomes, including: prevalence of specific diseases, overall mortality, and summary measures of population health including life expectancy and disease-free life expectancy. The term "dynamic" is used to describe how it uses actual population data and, hence, accounts for changing population compositions, risk factor prevalences, and disease burden in each country.

At its core, DYNAMO-HIA simulates a population (birth, death, and zero migration) and projects its future exposure to a risk factor. By applying the corresponding relative risks for disease incidence and mortality given individual risk factor and health status, future prevalences of chronic diseases and corresponding mortalities are calculated. This Markov-type approach accounts for competing mortality risk, i.e. despite multiple risks of dying an individual only dies once. Using a dynamic model allows us to assess the development of population health over time, i.e. in every additional year projected, mortality and disease incidence and the resulting disease prevalences are calculated for the entire population, taking account of age, sex, and risk factor prevalence. Such a dynamic approach recognizes how a) changes at the individual level take time to become visible at the population level, and that b) population composition is constantly changing. [113]

To specify a scenario, DYNAMO-HIA requires the initial risk factor prevalence, the new prevalences resulting from a price increase, and information on the future development of the risk factor exposure. For every new year, DYNAMO-HIA applies to each age- and sex-group the probability that individuals will stay in this risk factor group or will move to another one, e.g. how many will remain abstainers and how many will become light or moderate drinkers in the next age and calendar year. Over time, the specification of these (future) transition probabilities influences greatly the development of the risk factor prevalence. Bearing in mind that future individual behavior is

always subject to uncertainty and considering the short time span of our projection (10 years) we use the option of net-transition probabilities provided by DYNAMO-HIA. Net-transition probabilities keep the age-specific risk factor prevalence constant and do not take into account any future cohort effects.

In total, for each country seven scenarios are calculated. The reference scenario keeps the current level of consumption constant, thus estimating the pattern of population health in 10 years from now in the absence of intervention. This reference scenario, referred to as "business-as-usual", is the benchmark with which to compare the five alternative price scenarios, i.e. price increases of 20%, 40%, 60%, 80%, and matching the Finnish price level. The sixth scenario, where abstainers remain abstinent but all others become light drinkers, serves as the upper boundary, quantifying the potential gains of an alcohol policy solely based on pricing.

Results

Table 6.3 shows the number of deaths postponed, i.e. the difference between population size in year 10 when migration is zero and fertility the same across all scenarios, for males and females. For all scenarios and for all countries, a price increase leads to a postponement of deaths for both genders, though the effect is larger among males. The number of deaths postponed ranges from approx. 100 for a small country like Ireland experiencing a price increase of 20% up to approx. 20,000 for a large country like Germany, should it adopt the Finnish scenario. For all countries combined, approx. 28,000 male and 15,000 female deaths could be postponed if a 40% price increase was adopted. In the Finnish scenario, this would increase to approx. 54,000 (male) and 26,000 (female) deaths postponed.

Table 6.4 shows the reduction in the number of individuals with the different diseases, i.e. observed at year 10 as compared to the reference scenario across all countries included. In all scenarios, the number of individuals with disease is reduced relative to the reference scenario, but for men the greatest decrease is for stroke and diabetes, whilst for women the prevalence of cancer is affected most. For instance, in the Finnish scenario, after 10 years, the number of men with stroke is reduced by 97,850, while those with diabetes are 65,850 lower and those with cancer are 62,000 lower. In women the corresponding figures are 19,100 (stoke), 23,450 (diabetes), and 27,000 (cancer), respectively.

Table 6.3: Deaths postponed* projected 10 years into the future.

a) Males

Absolute Difference to Reference Scenario (Rounded)

	Denmark	Finland	France	Germany	Ireland	Italy	Netherlands	Poland	Spain	Sweden	UK	All Countries
20% price increase	250	150	1,650	1,950	100	1,800	350	750	1,050	150	4,150	12,300
40% price increase	550	350	3,800	4,450	200	4,150	750	1,750	2,350	350	9,400	28,100
60% price increase	900	550	6,550	7,650	350	7,150	1,350	3,050	4,050	550	16,200	48,400
80% price increase	1,350	850	8,800	11,300	450	10,100	1,850	4,500	5,900	850	24,150	70,100
Finnish Scenario	300	0	8,650	12,900	0	5,700	1,650	5,400	8,200	150	10,850	53,850
Most Light Drinkers	2,350	1,600	13,800	19,650	850	16,950	3,150	8,400	10,200	1,550	48,700	127,150

Difference as Percentage of the Difference between Reference Scenario and Most Light Drinkers Scenario

	Denmark	Finland	France	Germany	Ireland	Italy	Netherlands	Poland	Spain	Sweden	UK	All Countries
20% price increase	10.60%	9.40%	12.00%	9.90%	11.80%	10.60%	11.10%	8.90%	10.30%	9.70%	8.50%	9.70%
40% price increase	23.40%	21.90%	27.50%	22.60%	23.50%	24.50%	23.80%	20.80%	23.00%	22.60%	19.30%	22.10%
60% price increase	38.30%	34.40%	47.50%	38.90%	41.20%	42.20%	42.90%	36.30%	39.70%	35.50%	33.30%	38.10%
80% price increase	57.40%	53.10%	63.80%	57.50%	52.90%	59.60%	58.70%	53.60%	57.80%	54.80%	49.60%	55.10%
Finnish Scenario	12.80%	0.00%	62.70%	65.60%	0.00%	33.60%	52.40%	64.30%	80.40%	9.70%	22.30%	42.40%
Most Light Drinkers	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

b) Females

Absolute Difference to Reference Scenario (Rounded)

	Denmark	Finland	France	Germany	Ireland	Italy	Netherlands	Poland	Spain	Sweden	UK	All Countries
20% price increase	200	50	600	1,300	0	1,050	350	50	300	50	2,950	6,900
40% price increase	400	100	1,300	2,850	50	2,300	800	100	600	150	6,450	15,100
60% price increase	700	150	2,150	4,850	50	3,950	1,300	200	1,050	300	11,100	25,800
80% price increase	1,000	250	3,100	6,900	100	5,800	1,950	300	1,550	400	16,100	37,500
Finnish Scenario	250	0	3,050	7,750	0	3,100	1,700	400	2,250	100	7,450	26,000
Most Light Drinkers	1,700	350	4,550	10,150	150	9,100	3,100	250	2,350	650	31,200	63,600

Difference as Percentage of the Difference between Reference Scenario and Most Light Drinkers Scenario

	Denmark	Finland	France	Germany	Ireland	Italy	Netherlands	Poland	Spain	Sweden	UK	All Countries
20% price increase	11.80%	14.30%	13.20%	12.80%	0.00%	11.50%	11.30%	20.00%	12.80%	7.70%	9.50%	10.80%
40% price increase	23.50%	28.60%	28.60%	28.10%	33.30%	25.30%	25.80%	40.00%	25.50%	23.10%	20.70%	23.70%
60% price increase	41.20%	42.90%	47.30%	47.80%	33.30%	43.40%	41.90%	80.00%	44.70%	46.20%	35.60%	40.60%
80% price increase	58.80%	71.40%	68.10%	68.00%	66.70%	63.70%	62.90%	120.00%	66.00%	61.50%	51.60%	59.00%
Finnish Scenario	14.70%	0.00%	67.00%	76.40%	0.00%	34.10%	54.80%	160.00%	95.70%	15.40%	23.90%	40.90%
Most Light Drinkers	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

Table 6.4: Reduction in cases of selected diseases in all countries as a result of increases in alcohol price compared to the reference scenario, projected 10 years into the future (by gender).

a)Absolute Difference to Reference Scenario (Rounded)

	Males				Females			
	Stroke	Diabetes	Cancers	IHD	Stroke	Diabetes	Cancers	IHD
20% price increase	14,650	20,450	12,450	8,650	7,050	4,600	6,800	400
40% price increase	33,600	46,850	28,800	19,750	15,800	10,350	15,150	1,100
60% price increase	57,850	80,400	49,500	34,250	27,100	17,700	25,800	1,900
80% price increase	84,850	114,600	71,400	51,450	38,950	25,700	37,900	2,350
Finnish Scenario	65,850	97,850	62,150	27,100	23,450	19,100	27,100	300
Most Light Drinkers	161,500	263,050	134,050	69,600	66,550	57,800	69,600	6,450

b) Difference as Percentage of the Difference between Reference Scenario and Most Light Drinkers Scenario

	Males				Females			
	Stroke	Diabetes	Cancers	IHD	Stroke	Diabetes	Cancers	IHD
20% price increase	9.10%	7.80%	9.30%	10.00%	10.60%	8.00%	9.80%	6.20%
40% price increase	20.80%	17.80%	21.50%	22.80%	23.70%	17.90%	21.80%	17.10%
60% price increase	35.80%	30.60%	36.90%	39.50%	40.70%	30.60%	37.10%	29.50%
80% price increase	52.50%	43.60%	53.30%	59.40%	58.50%	44.50%	54.50%	36.40%
Finnish Scenario	40.80%	37.20%	46.40%	38.90%	35.20%	33.00%	38.90%	4.70%
Most Light Drinkers	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

The "Most-Light-Drinkers" scenario, i.e. where the share of abstainers in the population remains unchanged but all drinkers consume less than 20g/d, serves as the theoretical upper bound of the effect of policy on the average amount of alcohol consumed. The potential gains in terms of death postponed are substantial: approx. 127,150 male and 63,300 female deaths over a ten-year period. The Finnish scenario already achieves more than 40% of the potential postponement of deaths that could be reached if all excessive drinking was curbed. Similarly, a large burden of chronic disease could be avoided, accounting for approx. 263,500 men and 57,800 women with diabetes and approx. 161,500 men and 66,550 women with strokes.

Discussion

Main findings

Using a dynamic population health modeling tool, we have been able to quantify, for the first time, the effect of increasing the price of alcohol in reducing the burden of a range of alcohol-attributable diseases in the EU. The dynamic nature of the model allowed individuals to transition through risk categories over time. This produced a robust estimate of the health impact of a limited range of alcohol-attributable harms. Even a modest increase of alcohol prices would lead to an overall reduction in the incidence of diseases and number of deaths, even allowing for the protective effect of alcohol against ischaemic heart disease at some ages. The most extreme scenario, involving harmonizing prices to the Finnish level, would reduce the burden of disease substantially. Although extremely conservative (see section on Limitations below), we estimate that the minimal health benefits range from approx. 100 fewer deaths over a 10-year period for a minimal 20% price increase in a small country, such as Ireland, to 20,000 fewer deaths in a large country, such as Germany, should it increase alcohol prices to those in Finland. As males carry the larger burden of alcohol-related disease, in this model, mortality, and morbidity improvements from alcohol price increases were greater for men than for women for all included conditions except breast cancer.

Limitations

All models are simplified versions of reality and this is no exception. The effects of alcohol on health are especially difficult to model, data on exposure,

and in particular drinking patterns, are limited, and the association between exposure and some important adverse health outcomes is uncertain, not least because it is influenced by context. Although our estimates of all-cause mortality are based on published relative risks associated with different levels of consumption, thereby including deaths from causes beyond the specific ones examined here, and in particular injuries and violence, we were not able to incorporate pattern of consumption. Nor were we able to include the effects on those who are harmed by others who have consumed alcohol. These include, among others, rape, assaults, domestic violence, child abuse, and injuries occurring during crimes, road traffic injuries, and sexually transmitted infections. Nor were we able to quantify certain conditions specifically associated with alcohol, including cirrhosis, certain other cancers, neuropsychological conditions, prenatal alcohol exposure, and acute alcohol poisoning.

In calculating the consequences of the price change, we used a mean price elasticity of demand, derived from a recent meta-analysis, for all age, sex, and exposure groups, which did not differentiate by type of beverage (e.g. beer, wine, liquor), mode of consumption (e.g. on-trade, off-trade) or country. Whilst estimates of elasticities are known to be affected by these variables,[114] we were unable to produce with confidence a formula that could be used to adapt the elasticities to each country or, ideally, to each sex and age group. Nor could we take into account cross-price elasticities, that is, the effect on consumption on one beverage as a result of price increases on another. In our calculations we assume an increase in the price of pure alcohol consumed in grams per day and do not differentiate by beverage type or mode of consumption. Similarly, we did not account for potential changes in individual behavior intended to evade the effect of alcohol price increases, such as alcohol smuggling or home brewing.

The pattern of drinking, e.g. binge drinking, is only captured indirectly as only health effects of average consumption are reflected. It is uncertain how a change in consumption caused by price increases might affect, for example, youth initiation or binge drinking patterns and the resulting harm indicators. We assumed a uniform distribution of consumption in each different consumption level in the absence of evidence of either a uniform or non-uniform distribution, which may have weakened the observed estimate of effect. Our use of consumption categories meant that we could only measured the health benefits arising from transitioning between categories. However, there may be substantial health gains for people who moved from the upper to lower

ends of a given category.

Given that most of these limitations tend to underestimate the effects of price increases, our estimates must be seen as the minimum differences in effects with different scenarios.

A further limitation of DYNAMO-HIA is the absence of data on historical trends on incidence and prevalence of the diseases included in the model, which would be needed to predict future trends in the absence of interventions. Trend-free data were used, estimated using DisMod II software, as a neutral option.[69] However, this is less important as the limitation will apply equally to all scenarios.

Conclusion

Even with our very cautious assumptions, the health gains in terms of death postponed and chronic diseases cases reduced when harmonizing prices to the Finnish level are substantial. Nevertheless, even more modest increases in alcohol prices, mediated through taxation across Europe, offer the scope to prevent many premature deaths and much morbidity from a range of chronic conditions, demonstrating the harmful effects of excessive alcohol consumption. While there is much more to be done to capture the full effects of alcohol on health, the use of a dynamic population projection model contributes to the current EU dialog and provides support for measures that achieve both a substantial increase in, and harmonization of, the minimum duty on alcohol.

8

Discussion and Conclusion

A tool which has many purposes is not usually as efficient for anyone of them as a specialized tool developed solely for a single purpose.
M. G. Kendall

Findings

A quantification tool that is adequate to function as a standard tool in HIA must satisfy at least six criteria. The first three criteria – real-life population, dynamic projection, explicit risk factor states – warrant that the technical core of the model sufficiently lives up to the standards of established scientific modeling practice. The last three criteria – modest data requirements, rich model output, generally accessible – ensure a wide usability of the tool accounting for the limited resources of a decision making process.

Such a standard tool did not exist as of 2008, the year in which the systematic review took place. Of the six models identified to be sufficiently generic, i.e. allowing for various and multiple diseases and different risk factors, none could be considered a potential standard tool. There is an evident gap between the advanced models that have no or limited general accessibility (such as POHEM and RIVM-CDM), and the (over-)simplifying but more accessible models (such as GBD, MSLT and partly PREVENT). This situation probably arises because none of the reviewed models (except for ARMADA) was initially intended to be a software application for wider public use for the (relatively recent) task of quantification in HIA. Furthermore, the systematic review demonstrates that any tool that intends to fill this gap in the future needs to put equal weight on appropriate simulation methodology as well as data requirements that can be widely met while also being end-user friendly.

DYNAMO-HIA synthesizes several methodological elements, satisfying the aforementioned criteria: A micro-simulation of exposure information with macro-simulation of diseases and survival. This allows generic simulation of detailed and complex individual risk-factor biographies while avoiding the need for large simulated populations due to the relative rareness of chronic disease events, reducing calculation time considerably. The challenge of keeping the input data requirements feasible is met by including parameter calculation routines which use marginal population data to estimate both the transitions between states and the initial state occupancy. By combining feasible data needs with general accessibility and user friendliness within the causal framework of HIA, DYNAMO-HIA is a potential standard tool for health impact assessment based on epidemiological evidence. As it is intended for a broad user base, it comes with a detailed documentation clearly describing practical application and the underlying model structure.

Quantifying the potential health gains and losses across risk factors, i.e. alcohol, BMI, and smoking, for several countries requires the definition of a sensible benchmark. No single-best approach exists. Hence, a policy-making perspective was chosen to select feasible, i.e. observed, "worst"- and "best"-practice risk factor prevalences. The relative comparison showed that smoking is still the risk factor with the largest potential for health gains and health losses across the analyzed eleven EU-countries, while BMI has a comparatively large effect on morbidity. Applying the best practice smoking prevalence would yield the largest gains in life expectancy with 0.4 years for males and 0.3 years for females, while the worst practice smoking prevalence would lead to the largest losses with 1.2 years for males and 1.4 years for females. In terms of morbidity, the results differ by gender. For males, the best practice smoking prevalence would increase the disease-free life years the highest with 0.4 life years, whereas for females the best practice BMI prevalence promises the largest gains with an additional 0.7 disease-free life years.

Increasing the common excise taxes on alcohol sales across all EU-countries, as it is currently debated, would lead to an overall reduction in consumption and in turn to a reduction in morbidity and mortality attributable to alcohol consumption. Harmonizing alcohol prices to the Finnish level could postpone approx. 54,000 male and approx. 26,000 female deaths and reduce the prevalence of a number of chronic diseases over the next 10 years. Moreover, curbing excessive drinking completely could lead to substantial population health gains.

Reducing either the prevalence of overweight/obesity at age 18 or reducing the lifelong probability of becoming overweight/obese has a detectable effect in reducing the overall BMI level of the cohort, and in turn reducing the prevalence of chronic diseases and improving life expectancy. But even a small increase in the transition probability of becoming overweight/obese throughout the life course would offset a substantial reduction in prevalence when entering adulthood. In the example cohort, reducing the prevalence of overweight/obese at age 18 by 100% could increase cohort life expectancy by 0.22 life years whereas a reduction in transition probabilities to become overweight/obese by 100% could increase cohort life expectancy by 0.51 life years.

Challenges in quantification using DYNAMO-HIA

In the second part of the this thesis the usefulness of DYNAMO-HIA was tested on different real-life policy applications. Conducting these applications generated several insights concerning methodology, data availability, effects of policy and, more generally, validity of quantification in health impact assessment.

Methodological considerations

DYNAMO-HIA's methodology is strongly founded in demographic theory, i.e. that population composition changes through (mortality-, fertility-, and migration-) transition rates.[113] Hence, a characteristic of DYNAMO-HIA as compared to a number of other population health models is the use of explicit risk factor states that allow the specification of transitions between different (health) states. This improves the formal accuracy of the model and avoids bias through mortality selection.[36] But it also increases plausibility as this approach does not fix a future prevalence which may require unrealistic transitions in the simulated population.

A minor limitation of DYNAMO-HIA is that in its current stage it does not allow the specification of a fertility rate but only the number of births per year and that it assumes a closed population, i.e. zero migration. Both limitations have been accepted deliberately as those elements require additional data but would affect the comparisons between scenarios for the same population in most cases merely slightly.[113] Similarly, DYNAMO-HIA does not allow

time-trends in the specification of transitions, e.g. an autonomous decline in mortality by 5% each calendar year. Again, this has been a deliberate choice as this would increase data demands while such an autonomous trend would affect all scenarios in a given comparison.

An important consideration in modeling is how statistical uncertainty in the input data is accounted for. The standard approach in health decision models is to conduct a probabilistic sensitivity analysis (PSA). A PSA is very demanding on both data requirements and computational resources and, hence, onerous to implement as a standard option. In its current form, however, DYNAMO-HIA can be run via batch mode enabling a competent user to conduct such a PSA.[119,120]

Data availability

The more accurate a model aims to mirror a real-life process (and by that improving formal validity), the more data is needed. But even at the national level, comprehensive health data is not readily available (see Table C.1 for details). For example, in the collection of risk factor exposure data on smoking, alcohol consumption, and BMI, the DYNAMO-HIA project was not able to locate such data for over a third of all EU-27 countries. Although the problem is more pronounced for new member states and smallers countries, it is still widely spread. In a number of cases, regional data or data from neighboring countries is used as proxies.

Furthermore, the quality of the available data is often problematic. For example, the disease data needed for DYNAMO-HIA – consisting of incidence, prevalence, and excess mortality – can be inconsistent when taken directly from the data sources. That means, the reported incidence data does not correspond to the reported levels of prevalence or excess mortality or vice versa. However, such inconsistencies, inter alia caused by shortcomings in the data collection process or through time trends in the data, must be corrected for to obtain reliable results. This can be done by employing the IPM framework as outlined in Appendix B, requiring often disease-specific expertise. This has been done for all the data prior to inclusion into DYNAMO-HIA, i.e. the included data is consistent.

For the purpose of HIA, which is not only interested in the overall effects but also their distribution within a population, more detailed data on (social-)inequality is needed. These are in particular risk factor exposure, the cor-

responding relative risks, and data on disease prevalence/incidence/excess mortality by socioeconomic status. The modeling approach taken by DYNAMO-HIA is in principle able to account for health developments by socioeconomic status, but the data to do so is often lacking.

Effects of policy on exposure often unknown

A recurring challenge in all applications of DYNAMO-HIA is to conceptualize and quantify the effect a policy has on the risk factor that is being modeled. In particular, how does the risk factor prevalence change within the population (who improves, who worsens, by how much), when does this change occur, and is this a lasting or a one-time effect?

Currently, only sporadic knowledge of effects of policies on health behavior exists. For selected policies directed towards alcohol and tobacco, economists do research into the subsequent change in consumption. Yet, often those results are not detailed enough to be applied when modeling a population stratified by age and sex. Identifying and quantifying such relationships will be an ongoing core task of HIA. This requires a multi-disciplinary approach as many relationships defy orthodox thinking.

But DYNAMO-HIA allows to go the other way: Quantifying health benefits of several risk factor prevalences and by that identifying those which are desirable; policy makers could then use a variety of policy tools at their disposal to direct a population towards these desirable risk factor prevalences.

Validity in Health Impact Assessment

A model is always a simplified version of reality, and in that sense a model should not be classified as right or wrong but judged as valid for the problem it is applied to. For the purpose of judging quantitative models for HIA, validity can be conceptually split into three elements: *formal validity*, *plausibility*, and *predictive validity*.[2]

The first element, *formal validity*, assesses whether correct methods have been applied correctly. DYNAMO-HIA is deliberately based on formal demographic theory of population change and uses standard epidemiological "dose-effect" relationships.[113,121] In doing so it utilizes established epidemiological theory to causally model the effect of risk factor exposure on chronic disease incidence and mortality.

The second element, *plausibility*, refers to the degree to which an observer deems the theoretical framework as understandable, applicable, and plausible. The design of the model structure of DYNAMO-HIA has been conducted in consultation with both modeling experts and HIA practitioners. Furthermore, the applications to real-life policy problems – in close collaboration with established experts for the risk factors question – demonstrate that DYNAMO-HIA makes a valuable and accepted addition to the particular research field.

The third element, *predictive validity*, is the degree to which predictions are confirmed by (future) facts. Prediction validity is the holy grail of social science and most difficult to establish, if at all. It is (prohibitively) time consuming to check the results of a prediction model. For example, the effect of an intervention on cohort life-expectancy would take some 100 years of waiting. Moreover, the counterfactual cannot be observed, i.e. what would the reality be if another course of action would have been taken. Lastly, the *Oedipus-effect* might materialize, i.e. knowing the future changes the behavior in such a way as to make the prediction invalid.[122]

What does DYNAMO-HIA add?

Currently, quantification in HIA is (too) seldom attempted and those studies that try to quantify lack uniformity in their approach.[2] Chapter 2 demonstrated that so far no publicly available potential standard tool for quantification in health impact assessment existed. This changed with the advent of DYNAMO-HIA, being designed for this purpose from the onset. However, quantitative HIA is a field of rapidly growing interest with new models being designed,[123] such as the Impact Calculation Tool (ICT) that is part of the INTARESE/HEIMTA project that solely focuses on environmental exposure.[124,125] Furthermore, software models are often subject to ongoing or sporadic changes. For example, PREVENT implemented a cohort option allowing to model inter alia smoking status explicitly for at least part of the population.[126]

Another major addition of DYNAMO-HIA is the accompanying, ready-to-use data set. This comparatively comprehensive data set covers all EU-27 countries. It contains data on three major life-style-related risk factors – alcohol consumption, BMI, and smoking – and causally connected major chronic diseases, such as IHD, several cancers, COPD, and diabetes. When some data

Table 8.1: Updated comparison of the reviewed models against the evaluation criteria.

	Criterion					
	Real-life population	Dynamic projection	Explicit risk factor states	Modest data requirements	Rich model output	Generally accessible
ARMADA	+	+	+	-	-	-
GBD	-	-	-	+	-	+*
POHEM	+	+	+	-	+	-
PREVENT	+	+	-	+	+	-
RIVM-CDM	+	+	+	-*	+	-
Proportional Multi-state Life Table (MSLT)	-	-	-	+	+	+*
DYNAMO-HIA	+	+	+	+	+	+

*with some restrictions

is missing for a particular country, it implies that – at least at the time of the data collection – this data virtually did not exist. This unique data set allows now for the first time a comparatively rapid assessment of the consequences of changes in risk factor exposure for many EU-countries, making DYNAMO-HIA truly a ready-to-use tool.

But DYNAMO-HIA is not a one-stop-tool. A quantification tool is only one element in a longer process that is needed to asses the impact of a policy or intervention on health. A lot of decisions have to be taken before and after the use of the tool that influence the results. And for some, more specialized applications – such as the health effects of the local dispersion patterns of fine particulate matter – more specialized tools are recommendable.[127] Nevertheless, DYNAMO-HIA is certainly a first-stop-tool, allowing the quantification of a large range of salient scenarios. And even more specialized research questions could be first analyzed with DYNAMO-HIA to gauge whether the possible magnitude of the effects warrant a more detailed analysis. Such a tool increases the uniformity and comparability of results.

Implications

Implications for modeling

The main bottleneck in population health modeling is not lack of more complex mathematical models, but rather scarcity and inconsistencies of the available health data.[128] Hence, a focus of modeling work should lie in improving our understanding and, subsequently, use of the available data. Tools such as DisMod II are aiming to extract more information from incomplete or inconsistent data.[69] Moreover, population health models themselves should aim to reduce the data needs without sacrificing accuracy, such as the parameter estimation module in DYNAMO-HIA. Lastly, the modeling of uncertainty of the input data is a potential area of improvement. In particular, population health models should aim to increase the ease of use for practitioners.

Implications for data collection

The data collection efforts revealed several shortcomings of the availability within the EU-27 countries. For a surprising number of member states, data on risk factor prevalence or diseases is only partly or not at all available. In some cases, data is only based on regional surveys or is already somewhat dated and, hence, not fully representative. Furthermore, some age groups are often under-represented or not covered at all and, additionally, disease data often showed inconsistencies affecting the overall quality. The responsible public health institutions should increase their efforts in the collection of timely and consistent data, at least for major risk factors – such as tobacco, alcohol, or BMI – and for chronic diseases that contribute substantially to the morbidity or mortality burden.

Another, more ambitious implication for the current practice of public health data collection is a stronger incorporation of data about social inequalities.[129] Epidemiology demonstrates increasingly that a number of health consequences due to risk factor exposure vary not only by age and sex but are also influenced by social gradient or life course exposure. Yet, even at the national level basic data is hard to get. This problem is even more pronounced for regional data, which might often be necessary as regions within a nation can vary substantially in their health profile for various reasons.

Implications for health impact assessment

The field of health impact assessment has profited greatly from epidemiological insights and evidence. The causal chain between risk factor exposure and individual health is for many applications already sufficiently strong. An area of research that is still underdeveloped is the connection between policies and changes in individual risk factor exposure. A comprehensive theory for this question is impossible to achieve as it always depends on a number of factors, such as the risk factors in question, the policy tool used, and the context – as many relationships between policy and risk factor exposure are not stable over space and time. Nevertheless, some aim to ameliorate this problem by providing guidelines on how to quantify the most salient policies, such as the *Tobacco Toolkit* of the World Bank, possibly leading to a library of standard policies. There is utmost need for an increase in evaluation studies, assessing the efficacy and effectiveness of standard policies and their applicability across different contexts. A promising avenue for this is the increased access to longitudinal data by social epidemiologists and sociologists. Those data sets should be scrutinized increasingly to identify factors that lead to a change in individual risk factor behavior. This should be done while also paying attention to what moderating variables, such as income or eduction, contribute.

Implications for policy

The three applications of DYNAMO-HIA yield implications for policy. Quantifying potential future health gains and losses due to life-style-related risk factors require an understanding of an applicable benchmark. Deriving such benchmarks is a non-trivial task and will require on-going discussion as a given approach always makes (implicitly or explicitly) a value judgment. The benchmark developed in this thesis shows that reducing smoking should still be high on the EU-policy agenda. But obesity will likely become increasingly important, calling for policies that are effective through the whole life course. Clearly, also an overall reduction in alcohol consumption will lead to significant health gains throughout the EU. The EU exhibits a wide variety of price levels across countries as measured by the consumer price index. A worthwhile goal is to harmonize existing alcohol prices to a higher level. But price policy can only be one aspect of alcohol policy as the context in which drinking takes place is of importance as well. The analysis of the life course devel-

opment of the BMI of a cohort demonstrates that potential health gains of an childhood focused eradication of overweight/obesity can be quickly offset by a small, constant increase in the lifelong risk of becoming overweight/obese.

Conclusion

DYNAMO-HIA, fulfilling all criteria for a potential standard tool for quantification and already being equipped with a unique data set, is an useful addition to the field of HIA. Selected applications demonstrate the applicability, plausibility, and usefulness of DYNAMO-HIA and its results. DYNAMO-HIA, being designed for this purpose from the onset, may improve the number of quantitative HIAs significantly.

9 Summary

English

This thesis originated from and contributed to the DYNAMO-HIA project. The aim of the DYNAMO-HIA project was to construct a publicly available, potential standard tool for health impact assessment (HIA), collect and make available the necessary EU-wide data for the standard tool, and apply this new tool to selected policy questions.

HIA is evaluation of policies, projects or proposals concerning their effects on human health. HIA differs from other approaches in its focus on all policies or proposals, i.e. it is not solely focused on health policies but on any policy area. An HIA exercise usually has three main objectives: First, to *predict* the impact of a policy, second, to allow *participation* of stakeholders in the assessment process, and, third, to *inform* the decision making process. Quantification of the policy options under discussion, i.e. the expected future health changes, are done in the stage of *effect analysis*. Within this stage, three distinctive tasks have to be addressed: 1) Description of the baseline situation, i.e. what is the future situation without the policy, 2) estimation of change in exposure to determinants of health, i.e. how does the policy affect risk factor exposure, and 3) estimation of change in health outcomes, i.e. how does the change in risk factor exposure affect population health.

Part I: Methodological Foundation

Currently, quantification in HIA is seldom undertaken due to the lack of a standard tool. A tool should aim for technical accuracy in the prediction of the effects of interventions on population health, and also be effective in the applied setting of an HIA. Hence, six criteria were deduced a potential stan-

dard tool must fulfill. The first three criteria (*real-life population*, *dynamic projection*, and *explicit risk factor states*) ensure that the model structure is sufficiently advanced to model changes in risk factor exposure over time in a real-life population in a transparent way. The last three criteria (*modest data requirements*, *rich model output*, and *generally accessible*) ensure a wide usability by accounting for the constraints of a decision-making process.

A systematic review showed, that of 2008, no existing model can serve as a standard tool for quantification in HIA. There is an evident gap between the advanced models that have no or limited general accessibility (such as POHEM and RIVM-CDM), and the (over-)simplifying but more accessible models (such as GBD, MSLT and partly PREVENT). This situation probably arises because none of the reviewed models (except for ARMADA) was initially intended to have a software application for wide public use for the (relatively recent) task of quantification in HIA.

Conventional Markov models for risk factor exposure on chronic disease potentially contain a large number of states (risk factor and disease combinations), providing a challenge both technically (keeping execution time and memory use down) and practically (estimating the model parameters and retaining transparency). Hence, for the DYNAMO-HIA model we proposed a combination of micro-simulation of exposure information with macro-simulation of diseases and survival. This allows simulation of exposure detail while avoiding the need for large simulated populations due to the relative rareness of chronic disease events. Further efficiency is gained by splitting the disease state space into smaller spaces, each containing a cluster of diseases that is independent of the other clusters. The challenge of feasible input data requirements is met by including parameter calculation routines that use marginal population data to estimate both the transitions between states and the initial state occupancy.

The DYNAMO-HIA model can simulate risk factor exposure in three different forms: *continuous*, in *classes* (up to 10 categories), and in classes where for one class the *duration of class membership* has health consequences. The model accommodates up to three different diseases: *chronic diseases*, *partly acutely fatal diseases*, and diseases where the excess mortality depends on the *duration of the diseases*. The policy-induced change in risk factor prevalence and/or risk factor transition rates will be determined by the user. Hence, the tool can be used after the user has specified the effect of policies on health determinants. Several population-based health *outcome measures* – such as life

expectancy or disability-adjusted life expectancy – are readily available to quantify the difference between the reference and the different policy scenarios.

The model is available with internally consistent incidence, prevalence, and disease mortality (IPM) data by age and sex for nine diseases (IHD, diabetes, stroke, COPD, breast-, colorectal-, lung-, oral-, and esophageal-cancer). Furthermore, data on three risk factor prevalences (BMI, smoking, alcohol consumption) and relative risks quantifying the association between the risks factors, the diseases, and total mortality are included. The collected data has been compiled from already existing data sources. The objective was to collect these data for each EU member state. However, availability of data varied tremendously but for almost 80% of the EU population complete coverage was possible.

Part II: Selected Applications

The first application focuses on a cross-country comparison of modifying the exposure to the life-style-related risk factors alcohol, BMI, and smoking in eleven EU-countries, covering approx. 80% of the EU-27 population. Effective policy-making requires quantification of realistic gains and losses in population health, but the population-level health effects of these three risk factors often depend on a range of population-specific characteristics. Applying the DYNAMO-HIA model, we dynamically projected for every country the effects of potential health gains and health losses using feasible, i.e. observed elsewhere, risk factor prevalence for three different risk factors as benchmarks. The effects of the "worst practice", "best practice", and the currently observed risk factor prevalence on population health are projected for 10 years into the future and changes in life expectancy, disease-free life years, disease prevalences, and cumulative mortality are reported. Applying the best practice smoking prevalence would yield the largest gains in life expectancy with 0.4 years for males and 0.3 years for females while the worst practice smoking prevalence would also lead to the largest losses with 1.2 years for males and 1.4 years for females. In terms of morbidity, the results differ by gender. For males the best practice smoking prevalence would increase the disease-free life years the highest with 0.4 life years whereas for females the best practice BMI prevalence promises the largest gains with an additional 0.7 disease-free life years. Smoking is still the risk factor with the

largest potential health gains and losses, with the exception of the effect of BMI female disease-free life expectancy.

The second application focuses on the high level of alcohol consumption in Western Europe and corresponding adverse health effects. There is clear evidence that increased alcohol prices reduce alcohol consumption. We quantified some of the potential health consequences of increasing price on alcohol by projecting varying tax rates for 10 years into the future. Using alcohol consumption data for eleven EU-countries, covering 80% of the EU-27 population, we projected the expected changes in selected measures of population health that might arise from changes in alcohol price. Even a modest price increase of 20% leads to fewer cases of stroke, diabetes, and cancer, and fewer deaths in both men and women. Effects are larger in men. An increase in alcohol prices towards those currently in Finland (the highest in the EU) would postpone approx. 54,000 male and approx. 26,000 female deaths. Moreover, the prevalence of a number of chronic diseases would be reduced, in men by approx. 97,850 individuals with stroke, 65,850 with diabetes and 62,000 with cancer, and in women by about 19,100 (stroke), 23,450 (diabetes), and 27,000 (cancer), respectively. Completely curbing excessive drinking throughout the EU would lead to substantial gains in population health, while harmonizing prices to the level seen in Finland achieving some of those gains.

The third application compares the long-term health consequences of two different strategies of reducing obesity, i.e. childhood interventions that reduce the share of overweight/obese when entering adulthood and policies focusing on the risk of becoming overweight or obese throughout the life course. Using the case of a cohort of English males followed throughout their life course, we simulated the effect of either a) reducing the share of overweight/obese when entering adulthood, b) reducing the probability of becoming overweight/obese throughout their adult life, and c) combinations of both interventions. By employing the dynamic population health model DYNAMO-HIA, we projected the development of the share of overweight/obese and the resulting health consequences in terms of prevalence of related major chronic diseases, disease-free life years, and cohort life expectancy. Changing either policies can have a lifelong effect in reducing the overall BMI level of the cohort, and, in turn, reducing the prevalence of chronic diseases and improving life expectancy. A given percentage reduction in transition probability of becoming overweight/obese and not changing the initial prevalence of overweight/obese has a larger magnitude on the lifelong re-

duction of obesity and, hence, on health than same percentage reduction of prevalence at age 18 only. Furthermore, even a small increase in the probability of becoming overweight/obese throughout the life course would offset a substantial reduction in prevalence of overweight/obesity when entering adulthood.

Conclusion

So far, quantification in HIA is (too) seldom attempted and those studies that try to quantify lack uniformity in their approach. DYNAMO-HIA, being designed for this purpose from the onset, may improve the number of quantitative HIAs significantly. The possible use of generic risk factors and the accompanying, ready-to-use data set allows now for the first time a comparatively rapid assessment of the consequences of changes in risk factor exposure for many EU-countries, making DYNAMO-HIA a ready-to-use tool. DYNAMO-HIA, fulfilling all criteria for a potential standard tool for quantification and already being equipped with a unique data set, is an useful addition to the field of HIA. Selected applications demonstrate the applicability, plausibility, and usefulness of DYNAMO-HIA and its results.

Dutch

Dit proefschrift is ontstaan uit en heeft bijgedragen aan het DYNAMO-HIA project. De doelen van het DYNAMO-HIA project waren het ontwikkelen van een publiek toegankelijk, potentieel standaard instrument voor gezondheidseffectschatting (GES, in het Engels "Health Impact Assessment"), het verzamelen en beschikbaar maken van de noodzakelijke EU-brede gegevens voor het standaard instrument, en het toepassen van dit instrument om geselecteerde beleidsvragen te beantwoorden. GES evalueert beleid, projecten of voorstellen met betrekking tot hun effecten op de gezondheid van de mens. GES verschilt van andere benaderingen omdat het niet alleen gericht is op gezondheidsbeleid maar op alle beleidsterreinen of voorstellen. Een GES exercitie heeft meestal drie hoofddoelstellingen: ten eerste, om de impact van het beleid te voorspellen, ten tweede om de participatie van de belanghebbenden bij het evaluatie proces mogelijk te maken, en ten derde om de besluitvorming te informeren. Het kwantificeren van de beleidsopties, dat wil zeggen de verwachte veranderingen in de gezondheidstoestand ten gevolge van

de beleidsoptie, wordt uitgevoerd in de effect analyse fase. Tijdens deze fase moeten drie verschillende taken uitgevoerd worden: 1) beschrijving van de nul situatie (toekomstige situatie zonder het beleid), 2) schatting van de verandering in de blootstelling aan determinanten van gezondheid, dat wil zeggen hoe het beleid de risicofactorblootstelling verandert, en 3) bepaling van veranderingen in gezondheid, dat wil zeggen hoe de verandering in de risicofactorblootstelling gezondheid van de bevolking beïnvloedt.

Deel I: Methodologische onderbouwing

Momenteel vindt kwantificering in GES zelden plaats omdat een standaard instrument ontbreekt. Een dergelijk instrument moet streven naar technische nauwkeurigheid bij het voorspellen van de effecten van interventies op de volksgezondheid. Verder moet het instrument ook bruikbaar zijn in de toegepaste setting van een GES. Hiervoor zijn zes criteria opgesteld waaraan een standaard instrument moet voldoen. De eerste drie criteria (gebruik van werkelijke bevolkingsaantallen, dynamische projectie en expliciete risicofactortoestanden) moeten ervoor zorgen dat de modelstructuur voldoende geavanceerd is om veranderingen in de risicofactorblootstelling op opeenvolgende tijdstippen in de bevolking op een transparante wijze te modelleren. De laatste drie criteria (bescheiden gegevensvereisten, rijke model uitvoer en algemene toegankelijkheid) zorgen voor een brede inzetbaarheid, rekeninghoudend met de aanwezige beperkingen binnen een besluitvormingsproces.

Een systematische review toonde aan dat per 2008 geen bestaand model kan dienen als een standaard instrument voor de kwantificering van GES. Er is een duidelijke kloof tussen de geavanceerde modellen die niet of beperkt algemeen toegankelijk zijn (zoals POHEM en RIVM-CDM), en de (over-) vereenvoudigde, maar beter toegankelijke modellen (zoals GBD, MSLT). Deze situatie ontstaat waarschijnlijk omdat van de onderzochte modellen (met uitzondering van ARMADA) geen enkel model aanvankelijk een softwareapplicatie heeft ontwikkeld die kan dienen voor een breed publiek om GES te kwantificeren.

Conventionele Markov modellen die de blootstelling van risicofactoren op chronische ziekte modelleren kunnen een groot aantal toestanden (risicofactor en ziekte combinaties) bevatten. Dit betekent zowel een uitdaging in een technisch opzicht (het beperken van de rekentijd en geheugengebruik)

als in een praktisch opzicht (het schatten van de modelparameters en het behoud van transparantie). Om deze reden hebben wij voor het DYNAMO-HIA model een combinatie van microsimulatie van risicofactor blootstelling met macrosimulatie van ziekten en overleving voorgesteld. Dit maakt het simuleren van de blootstelling aan een risicofactor in detail mogelijk, terwijl het de noodzaak vermijdt grote gesimuleerde bevolkingen te gebruiken die nodig zouden zijn gegeven de relatieve zeldzaamheid van chronische ziekte gebeurtenissen. Verdere efficiëntie wordt verkregen door het splitsen van de ruimte met ziektetoestanden in kleinere ruimtes, die elk een cluster van ziekten bevatten die onafhankelijk is van de andere clusters. Aan de uitdaging van beschikbare input data wordt voldaan door het opnemen van parameter schattingsroutines waarin de marginale bevolking wordt gebruikt om zowel de overgangen tussen toestanden en de verdeling van de initiële toestanden te schatten.

In het DYNAMO-HIA instrument kan risicofactorblootstelling op drie verschillende wijzen worden opgenomen: continue, in klassen (tot 10 categorieën), en in klassen met voor één van die klassen de duur van het verblijf in die klasse. Het model biedt plaats aan maximaal drie verschillende ziekten: chronische ziekten, deels acuut dodelijke ziekten, en ziekten waarvoor de oversterfte afhankelijk is van de duur van de ziekten. Beleidgeïnduceerde veranderingen in risicofactor prevalentie en/of risicofactortransities worden bepaald door de gebruiker van het model. Daarom kan de tool gebruikt worden nadat de gebruiker heeft opgegeven wat het effect is van beleidsmaatregelen op de risicofactor. Verschillende volksgezondheid uitkomstmaten - zoals levensverwachting of "disability-adjusted" levensverwachting - zijn direct beschikbaar om het verschil tussen de referentie en de verschillende beleidsscenario's te kwantificeren.

Het model bevat intern consistente incidentie, prevalentie en sterfte (IPM) gegevens naar leeftijd en geslacht voor negen ziekten (IHD, diabetes, beroerte, COPD, borst-, colorectum-, long-, slokdarm- en mond kanker). Bovendien bevat het model gegevens over de blootstelling aan drie risicofactoren (BMI, roken en alcoholgebruik) en relatieve risico's die de relatie tussen de risicofactoren, incidentie van ziekten en totale mortaliteit kwantificeren. De verzamelde gegevens zijn ontleend aan bestaande gegevensbronnen. Het doel was om deze gegevens voor elke EU-lidstaat te verzamelen, echter, de beschikbaarheid van gegevens varieerde enorm. Voor bijna 80% van de EU-bevolking was een volledige dekking mogelijk.

Deel II: geselecteerde toepassingen

De eerste toepassing richt zich op een internationale vergelijking van gezondheidseffecten die gepaard gaan met het wijzigen van de blootstelling aan de leefstijl gerelateerde risicofactoren alcohol, BMI en roken in elf EU-landen (samen ca. 80% van de EU-27 bevolking). Effectief beleid vraagt om kwantificering van realistische winsten en verliezen in de gezondheid van de bevolking, maar de effecten op de gezondheid van deze drie risicofactoren zijn vaak afhankelijk van een aantal bevolkingsspecifieke kenmerken. Met behulp van DYNAMO-HIA, hebben we voor elk land de gezondheidswinsten en verliezen van mogelijke, dat wil zeggen elders waargenomen, risicofactor prevalentie dynamisch geprojecteerd. De effecten van de "worst practice", "best practice", en de momenteel waargenomen risicofactor prevalentie op de volksgezondheid zijn geprojecteerd voor 10 jaar in de toekomst. Tevens zijn de veranderingen in de levensverwachting, ziektevrije levensjaren, ziekte prevalentie en de cumulatieve sterfte gerapporteerd. Het toepassen van de "best practice" prevalentie van roken zou de grootste winst in levensverwachting met 0,4 jaar voor mannen en 0,3 jaar voor vrouwen opleveren, terwijl de "worst practice" rook prevalentie ook zou leiden tot de grootste verliezen van 1,2 jaar voor mannen en 1,4 jaar voor vrouwen. In termen van morbiditeit verschillen de resultaten naar geslacht. De "best practice" prevalentie van roken zou voor mannen ook de hoogste toename van de ziektevrije levensverwachting betekenen (met 0,4 jaar), terwijl voor vrouwen de "best practice" BMI prevalentie tot de grootste winst met een extra 0,7 ziektevrije levensjaren leidt. Roken is nog steeds de risicofactor met de grootste mogelijke gezondheidswinsten en verliezen, met uitzondering van het effect van BMI op de ziektevrije levensverwachting van vrouwen.

De tweede toepassing richt zich op overmatig alcoholgebruik in West-Europa en de bijbehorende negatieve effecten op de gezondheid. Er zijn duidelijke aanwijzingen dat een toename van alcohol prijzen het alcoholgebruik verlaagt. We hebben een aantal van de mogelijke gevolgen van prijsstijgingen voor de gezondheid gekwantificeerd door de gezondheidseffecten van verschillende belastingtarieven voor 10 jaar in de toekomst te projecteren. Met behulp van alcoholgebruik gegevens van elf EU-landen, die 80% van de EU-27 bevolking omvatten, hebben we de verwachte gezondheidsveranderingen geprojecteerd die kunnen voortvloeien uit wijzigingen in de alcohol prijs. Zelfs een bescheiden prijsstijging van 20% leidt tot minder gevallen van

beroertes, diabetes en kanker, en het leidt tot minder sterfte bij zowel mannen als vrouwen. De effecten zijn groter bij mannen. Een stijging van de prijzen van alcohol gelijkstaand aan de prijzen van Finland (de hoogste in de EU) leidt tot het uitstellen van ca. 54.000 sterftevallen onder mannen en ca. 26.000 sterftevallen onder vrouwen. Bovendien zou de prevalentie van een aantal chronische ziekten worden verlaagd, bij mannen met ca. 97.850 gevallen voor beroerte, 65.850 voor diabetes en 62.000 voor kanker, bij vrouwen met ongeveer 19.100, 23.450 en 27.000, respectievelijk. Volledige terugdringen van overmatig drinken in de hele EU zou leiden tot een substantiële winst in gezondheid van de bevolking. Met de harmonisatie van de prijzen op het niveau van Finland zou een deel van deze winsten bereikt kunnen worden.

De derde toepassing vergelijkt de lange termijn gezondheidseffecten van twee verschillende strategieën die overgewicht proberen te verminderen, namelijk: interventies onder kinderen en adolescenten die het aandeel van overgewichtobesitas bij het bereiken van volwassen leeftijd verminderen en het beleid gericht op het verminderen van het risico op overgewicht of obesitas gedurende het volwassen leven. Gebruikmakend van gegevens van een cohort van Engelse mannen hebben we de volgende effecten gesimuleerd: a) de vermindering van het aandeel van overgewicht/obesitas op het moment van bereiken van de volwassen leeftijd bij beide interventies, b) de vermindering van de kans op overgewicht/obesitas gedurende hun volwassen leven bij beide interventies, en c) de effecten bij een combinatie van beide interventies. Door gebruik te maken van het dynamische volksgezondheid model DYNAMO-HIA, hebben wij de ontwikkeling van het aandeel van overgewicht/obesitas en de daaruit voortvloeiende gevolgen voor de gezondheid in termen van de prevalentie van gerelateerde ernstige chronische ziekten, ziektevrij levensjaren, en de levensverwachting geprojecteerd. Alle typen interventies hebben een levenslang effect, ze verminderen het totale BMI niveau van het cohort, en daardoor verminderen ze de prevalentie van chronische ziekten en verbeteren ze de levensverwachting. Een bepaalde procentuele vermindering in de kans op overgewicht/obesitas gedurende het volwassen leven (zonder verandering in de prevalentie van overgewicht/obesitas op het bereiken van de volwassen leeftijd) heeft een groter effect op de levenslange vermindering van obesitas, en dus op de gezondheid, dan eenzelfde procentuele vermindering van de prevalentie op 18 jarige leeftijd. Bovendien zou zelfs een kleine toename in de kans op overgewicht/obesitas gedurende het volwassen leven een aanzienlijke vermindering van de prevalentie van over-

gewicht/obesitas op het bereiken van de volwassen leeftijd teniet doen.

Conclusie

Tot nu toe heeft kwantificering in GES (te) weinig plaatsgevonden en de studies die dit wel hebben getracht missen een uniforme aanpak. DYNAMO-HIA, dat van begin af aan was ontworpen voor dit doel, kan het aantal kwantitatieve GES aanzienlijk doen toenemen. Generieke risicofactoren en de bijbehorende kant-en-klare dataset maken nu voor het eerst een relatief snelle beoordeling van de gevolgen van veranderingen in de blootstelling aan risicofactoren voor veel EU-landen mogelijk. Dit maakt DYNAMO-HIA een kant-en-klare gebruiksvriendelijke tool. DYNAMO-HIA, dat voldoet aan alle criteria voor een mogelijk standaard instrument voor kwantificering en reeds uitgerust met een unieke dataset, is een waardevolle aanvulling op het gebied van GES. Geselecteerde toepassingen tonen de toepasbaarheid, plausibiliteit, en het nut van DYNAMO-HIA en de resultaten ervan aan.

Appendix

A

DYNAMO-HIA project

This thesis originated from and contributed to the DYNAMO-HIA project. The DYNAMO-HIA project was an international research project 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 *Erasmus Medical Center Rotterdam (the Netherlands)*, *Institute of Public Health and the Environment (the Netherlands)*, *Catalan Institute of Oncology*, *International Obesity Task Force*, *London School for Hygiene and Tropical Medicine*, *Haughton Institute in Dublin*, and *Instituto Tumori in Milan (Italy)*.

The aim of the DYNAMO-HIA project was to develop and build an instrument to quantify the health impact of changes in health determinants as a result of different policies and apply it to selected life-style related health determinants and resulting diseases across EU-countries. The research project had three specific objectives. First, to develop and implement a stand-alone software tool (DYNAMO-HIA) to estimate the health impact of policies by comparing the population health impact of one or more policy interventions with a baseline scenario. Second, to compile and make publicly available data sets (consistent across EU-countries) on selected health determinants/risk factors (smoking, obesity, and alcohol consumption) and their effects on selected diseases. Third, to illustrate the tool by assessing the health effects of several health-relevant policy options with regard to these health determinants.

The DYNAMO-HIA project was structured into eleven work packages:

1. Coordination of Project: *Wilma Nusselder*, *Johan Mackenbach*
2. Dissemination of the Results: *Jet Smit*, *Lea den Broeder*

3. Evaluation of the Project: *Wilma Nusselder*
4. Model Specification: *Wilma Nusselder*, *Stefan K. Lhachimi*, *Hendriek C. Boshuizen*, *Pieter van Baal*
5. Construction of Software Tool: *Hendriek C. Boshuizen*, *Stefan K. Lhachimi*, *Rene Mondeel*, *Jan de Bruin*
6. Smoking: *Estevez Fernandez*, *Jose Maria Martinez-Sanchez*, *Esther Carabasa*
7. Overweight/Obesity: *Tim Lobstein*, *Rachel Jackson-Leach*
8. Alcohol: *Martin McKee*, *Joceline Pomerleau*, *Kate Charlesworth*
9. CVD and diabetes: *Kathleen Bennett*, *Tom O'Hara*
10. Cancer: *Andrea Micheli*, *Paolo Baili*
11. Definition of Scenarios: *Wilma Nusselder*, *Johan Mackenbach*, *Stefan K. Lhachimi*, *Margarete Kulik*, *Hendriek C. Boshuizen*

I have to thank all those who collaborated within the DYNAMO-HIA project, making it such a success. Without this joint effort, this thesis could not have been finished in the current form. In addition, I would like to extend my thanks to all those who have provided suggestions to the model specification during the design phase (at the DYNAMO-HIA workshop on 23rd of May, 2008, and by personal communication).

B

IPM-Framework for chronic disease data

Epidemiological data is often plagued with incompleteness and inconsistency. An *incidence-prevalence-mortality* (IPM) framework utilizes the (mathematical) connection between related disease variables. This allows to check for inconsistencies and/or to estimate missing data. Within the IPM-framework the widely used software DisMod II can assist discovering and correcting for such flaws. However, expert knowledge of the disease in question is still indispensable. DisMod II is a stand-alone Windows program that is publicly available. Originally, it has been developed for the Global Burden of Disease (GBD) project and is currently widely in use.[69,93]

DisMod II assumes a single-disease model (see Figure B.1). An individual is in either of four different states: healthy, diseased, dead from the disease, or dead from all-other causes. The main assumption is that *all-other-cause mortality* (m) is equal for diseased (D) and healthy (H), i.e. an additive mortality model. Furthermore, the transition rates between these states are assumed to be constant. In such a closed system, the figures in question just become an exercise in book keeping. Therefore, the application of an IPM-model is rather straightforward: the missing variable(s) can be estimated using the knowledge of the other variables. In the case of inconsistencies things are more difficult. Several reasons are conceivable for flawed epidemiological data, e.g. figures are from different regional contexts or have been measured differently (see for further discussion Kruijshaar et al.[93]). Another reason for inconsistencies is a possible violation of the steady state assumption. As prevalence is a stock variable and incidence is a flow variable, the basic IPM-model assumes that all rates are constant over time. This is often unlikely as incidence but

also mortality data often follow a time trend. An IPM-based software cannot automatically recognize and correct such errors. For this expert knowledge about the disease and the country in question is necessary. However, a software assists in identifying and diagnosing potential problems in epidemiological data.

Another requirement for input data in disease models is smoothness. Smoothness refers to the application of statistical procedures to transform the existing data into a shape that is less "rough". In an extreme case, smoothing would mean that the existing data is transformed into a single, flat line. Of course, this would mean that almost all information in the data has been "smoothed away" and this is not intended. But it shows clearly there is a trade-off between having "rough" and "spiky" raw-data versus an (over-)smoothed version of it. In the case of DYNAMO-HIA, the rationale for smoothing the input data is twofold. First, raw data exhibits often a lot of spikes, jumps, or outlying data which is a sign of inconsistencies, measurement errors, unusual observations, or just a result of (too) small sample sizes. Second, for modeling within DYNAMO-HIA these abrupt changes can distort the results of the calculation, leading to even larger and, hence, even more implausible jumps in the output than in the input.

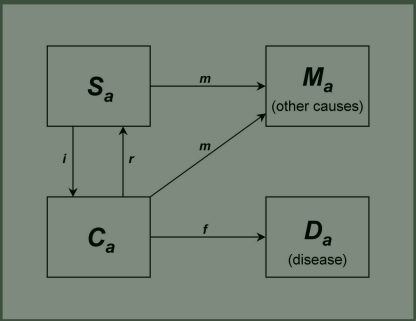

Figure B.1: The conceptual disease model of the IPM-Framework:[69] S:= number of healthy people, i.e. without the disease under consideration, C:= number of diseased people, D:= number of people dead from the disease, and M:= number of people dead from all other causes, with the subscript $_a$ denotes the age group. There are four transition hazards: i:= incidence, r:= remission, f:= case fatality, and m:= all other mortality.

C

Data overview

The software tool DYNAMO-HIA is equipped with an extensive data set of disease data, risk factor data, and the corresponding relative risks by gender and single year of age. This chapter displays the data (for simplification aggregated into age groups) collected within the DYNAMO-HIA project by the involved work packages (see Appendix A for details). The following Table C.1 shows for which of the EU-27 countries data was available. The design of the tool allows to enter further data for other countries and/or diseases and risk-factors. More details, in particular concerning the sources of the raw data and further adjustment procedures, are given in the work package documentation (available at www.DYNAMO-hia.eu).

Table C.1: Overview of available population-level data in DYNAMO-HIA (Y=Yes, N=No)

	Incidence, prevalence, and excess mortality (IPM) data					Risk factor exposure data		
	Diabetes	IHD	Stroke	COPD	Cancers	Obesity	Alcohol	Smoking
Austria	N	N	Y	N	Y	Y	N	N
Belgium	N	N	Y	Y	Y	Y	N	Y*
Bulgaria	N	N	N	N	Y	N	N	N
Cyprus	N	N	Y	N	N	N	N	N
Czech Republic	N	N	Y	Y	Y	Y	Y	Y*
Denmark	Y	Y	Y	Y	Y	Y	Y	Y*
Estonia	N	N	Y	Y	Y	N	Y	Y*
Finland	Y	Y	Y	Y	Y	Y	Y	Y*
France	Y	Y	Y	Y	Y	Y	Y	Y
Germany	Y	Y	Y	Y	Y	Y	Y	Y
Greece	N	N	Y	Y	N	N	Y	N
Hungary	N	N	N	Y	N	N	N	Y
Ireland	Y	Y	Y	Y	Y	Y	Y	Y*
Italy	Y	Y	Y	Y	Y	Y	Y	Y
Latvia	N	N	Y	Y	Y	N	Y	Y*
Lithuania	N	N	Y	Y	Y	N	N	Y*
Luxembourg	N	N	Y	N	N	N	N	N
Malta	N	N	Y	N	Y	N	Y	N
Netherlands	Y	Y	Y	Y	Y	Y	Y	Y
Poland	Y	Y	Y	Y	Y	Y	Y	N
Portugal	N	N	Y	Y	Y	Y	N	Y
Romania	N	N	N	N	N	N	N	N
Slovakia	N	N	N	Y	Y	N	N	Y*
Slovenia	N	N	Y	N	Y	N	N	N
Spain	Y	Y	Y	Y	Y	Y	Y	Y
Sweden	Y	Y	Y	Y	Y	Y	Y	Y*
United Kingdom	Y	Y	Y	Y	Y	Y	Y	Y

*No information on time since quitting

C.1 Disease Data

The DYNAMO-HIA project aimed to cover nine major chronic diseases: ischemic heart diseases (IHD), chronic obstructive pulmonary disease (COPD), stroke, diabetes, breast-, lung- colorectal-, esophageal-, and oral-cancer. The following Table C.2, shows the disease prevalence for each covered country by sex aggregated into six age groups.

Table C.2: Overview of disease prevalences by country used in DYNAMO-HIA aggregated by age (in percent). [130–132]

			Breast Cancer	Colo. Cancer	COPD	Diab.	Eso. Cancer	IHD	Lung Cancer	Oral Cancer	Stroke
Austria	0-15	M	0.0	0.0	n/a	n/a	0.0	n/a	0.0	0.0	0.0
	16-30		0.0	0.0	n/a	n/a	0.0	n/a	0.0	0.0	0.0
	31-45		0.0	0.0	n/a	n/a	0.0	n/a	0.0	0.0	0.1
	46-60		0.0	0.4	n/a	n/a	0.0	n/a	0.1	0.2	1.1
	61-75		0.0	1.7	n/a	n/a	0.0	n/a	0.5	0.5	5.1
	76+		0.0	3.9	n/a	n/a	0.1	n/a	0.8	0.8	13.8
	0-15	F	0.0	0.0	n/a	n/a	0.0	n/a	0.0	0.0	0.0
	16-30		0.0	0.0	n/a	n/a	0.0	n/a	0.0	0.0	0.0
	31-45		0.3	0.0	n/a	n/a	0.0	n/a	0.0	0.0	0.1
	46-60		1.8	0.3	n/a	n/a	0.0	n/a	0.1	0.1	0.6
	61-75		4.2	1.0	n/a	n/a	0.0	n/a	0.2	0.1	3.4
	76+		6.4	2.5	n/a	n/a	0.0	n/a	0.3	0.2	11.0
Belgium	0-15	M	0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0
	31-45		0.0	0.0	0.1	n/a	0.0	n/a	0.0	0.0	0.1
	46-60		0.0	0.3	0.9	n/a	0.0	n/a	0.1	0.2	1.1
	61-75		0.0	1.7	4.1	n/a	0.1	n/a	0.5	0.5	5.1
	76+		0.0	3.9	5.4	n/a	0.1	n/a	0.8	0.8	12.2
	0-15	F	0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0
	31-45		0.6	0.0	0.1	n/a	0.0	n/a	0.0	0.0	0.1
	46-60		2.9	0.3	1.1	n/a	0.0	n/a	0.0	0.1	0.7
	61-75		6.5	1.2	3.6	n/a	0.0	n/a	0.1	0.1	3.6
	76+		8.8	2.5	3.4	n/a	0.0	n/a	0.1	0.2	10.1
Bulgaria	0-15	M	0.0	0.0	n/a	n/a	0.0	n/a	0.0	0.0	n/a
	16-30		0.0	0.0	n/a	n/a	0.0	n/a	0.0	0.0	n/a
	31-45		0.0	0.1	n/a	n/a	0.0	n/a	0.0	0.0	n/a
	46-60		0.0	0.3	n/a	n/a	0.0	n/a	0.2	0.2	n/a

Continued on next page...

			Breast Cancer	Colo. Cancer	COPD	Diab.	Eso. Cancer	IHD	Lung Cancer	Oral Cancer	Stroke
	61-75		0.0	1.2	n/a	n/a	0.1	n/a	0.8	0.6	n/a
	76+		0.0	2.4	n/a	n/a	0.1	n/a	1.2	0.9	n/a
	0-15	F	0.0	0.0	n/a	n/a	0.0	n/a	0.0	0.0	n/a
	16-30		0.0	0.0	n/a	n/a	0.0	n/a	0.0	0.0	n/a
	31-45		0.3	0.0	n/a	n/a	0.0	n/a	0.0	0.0	n/a
	46-60		1.4	0.2	n/a	n/a	0.0	n/a	0.0	0.0	n/a
	61-75		3.0	0.8	n/a	n/a	0.0	n/a	0.1	0.1	n/a
	76+		4.1	1.5	n/a	n/a	0.0	n/a	0.1	0.2	n/a
Cyprus	0-15	M	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	0.0
	16-30		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	0.0
	31-45		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	0.1
	46-60		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	1.2
	61-75		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	5.1
	76+		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	12.9
	0-15	F	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	0.0
	16-30		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	0.0
	31-45		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	0.1
	46-60		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	0.7
	61-75		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	3.5
	76+		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	9.5
Czech Rep.	0-15	M	0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0
	31-45		0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.1
	46-60		0.0	0.4	1.1	n/a	0.0	n/a	0.2	0.2	1.9
	61-75		0.0	2.0	3.1	n/a	0.0	n/a	0.5	0.4	9.9
	76+		0.0	4.2	3.1	n/a	0.1	n/a	0.8	0.6	27.9
	0-15	F	0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0
	31-45		0.2	0.0	0.1	n/a	0.0	n/a	0.0	0.0	0.1
	46-60		1.4	0.3	1.1	n/a	0.0	n/a	0.0	0.1	0.9
	61-75		3.2	1.2	2.7	n/a	0.0	n/a	0.1	0.1	6.8
	76+		4.8	2.3	1.4	n/a	0.0	n/a	0.2	0.2	21.9
Denmark	0-15	M	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	31-45		0.0	0.0	0.1	1.7	0.0	0.2	0.0	0.0	0.2
	46-60		0.0	0.3	1.3	6.7	0.0	2.9	0.1	0.2	1.6
	61-75		0.0	1.3	4.8	13.4	0.0	10.1	0.3	0.6	5.6
	76+		0.0	3.3	6.1	15.3	0.1	18.4	0.6	0.9	13.2
	0-15	F	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
	31-45		0.3	0.0	0.1	2.3	0.0	0.1	0.0	0.0	0.1

Continued on next page...

			Breast Cancer	Colo. Cancer	COPD	Diab.	Eso. Cancer	IHD	Lung Cancer	Oral Cancer	Stroke
	46-60		2.1	0.3	1.4	5.0	0.0	1.4	0.1	0.1	0.8
	61-75		5.0	1.1	4.7	10.0	0.0	6.0	0.3	0.2	3.3
	76+		7.6	2.5	4.5	13.7	0.0	13.4	0.4	0.4	9.8
Estonia	0-15	M	0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0
	31-45		0.0	0.0	0.1	n/a	0.0	n/a	0.0	0.0	0.2
	46-60		0.0	0.2	0.9	n/a	0.0	n/a	0.2	0.1	2.2
	61-75		0.0	1.0	3.0	n/a	0.1	n/a	0.7	0.3	10.1
	76+		0.0	2.2	4.2	n/a	0.1	n/a	0.9	0.4	22.2
	0-15	F	0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0
	31-45		0.2	0.0	0.1	n/a	0.0	n/a	0.0	0.0	0.1
	46-60		1.1	0.2	0.9	n/a	0.0	n/a	0.0	0.0	1.1
	61-75		2.3	0.9	2.4	n/a	0.0	n/a	0.1	0.1	6.7
	76+		3.2	1.7	2.2	n/a	0.0	n/a	0.2	0.2	18.8
Finland	0-15	M	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	31-45		0.0	0.0	0.1	0.6	0.0	0.4	0.0	0.0	0.2
	46-60		0.0	0.2	1.1	4.1	0.0	5.2	0.1	0.1	1.5
	61-75		0.0	0.9	4.0	10.2	0.0	19.5	0.4	0.4	6.2
	76+		0.0	2.5	5.7	12.3	0.1	36.9	0.9	0.8	15.0
	0-15	F	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	31-45		0.3	0.1	0.1	0.4	0.0	0.2	0.0	0.0	0.1
	46-60		2.1	0.3	1.0	2.2	0.0	2.6	0.0	0.1	0.7
	61-75		5.4	0.9	3.2	7.3	0.0	12.0	0.1	0.2	3.8
	76+		8.0	2.1	3.6	11.9	0.0	27.0	0.3	0.4	11.7
France	0-15	M	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	31-45		0.0	0.0	0.0	1.2	0.0	0.1	0.0	0.1	0.1
	46-60		0.0	0.4	1.0	6.4	0.0	1.5	0.1	0.6	1.1
	61-75		0.0	1.8	4.0	15.2	0.2	5.9	0.4	1.6	4.3
	76+		0.0	4.3	5.2	16.2	0.4	10.5	0.6	2.4	10.1
	0-15	F	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	31-45		0.4	0.0	0.1	1.0	0.0	0.0	0.0	0.0	0.1
	46-60		2.6	0.3	0.9	4.1	0.0	0.6	0.0	0.1	0.4
	61-75		6.3	1.1	2.6	9.5	0.0	3.1	0.1	0.3	2.4
	76+		8.8	2.6	2.4	11.0	0.0	6.7	0.1	0.4	7.2
Germany	0-15	M	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Continued on next page…

			Breast Cancer	Colo. Cancer	COPD	Diab.	Eso. Cancer	IHD	Lung Cancer	Oral Cancer	Stroke
	16-30		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	31-45		0.0	0.0	0.1	1.3	0.0	0.3	0.0	0.0	0.1
	46-60		0.0	0.4	1.0	5.6	0.0	3.6	0.1	0.3	1.0
	61-75		0.0	1.8	3.8	12.6	0.1	14.5	0.7	0.7	5.0
	76+		0.0	4.2	5.5	13.6	0.2	27.1	1.4	1.0	13.8
	0-15	F	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	31-45		0.3	0.0	0.1	1.3	0.0	0.1	0.0	0.0	0.1
	46-60		1.8	0.3	0.9	5.7	0.0	1.7	0.0	0.1	0.5
	61-75		4.4	1.1	2.5	12.7	0.0	8.4	0.1	0.2	3.0
	76+		6.5	2.8	2.4	13.2	0.0	19.1	0.2	0.3	10.7
Greece	0-15	M	n/a	n/a	0.0	n/a	n/a	n/a	n/a	n/a	0.0
	16-30		n/a	n/a	0.0	n/a	n/a	n/a	n/a	n/a	0.0
	31-45		n/a	n/a	0.2	n/a	n/a	n/a	n/a	n/a	0.1
	46-60		n/a	n/a	1.1	n/a	n/a	n/a	n/a	n/a	1.7
	61-75		n/a	n/a	5.6	n/a	n/a	n/a	n/a	n/a	9.4
	76+		n/a	n/a	9.2	n/a	n/a	n/a	n/a	n/a	27.1
	0-15	F	n/a	n/a	0.0	n/a	n/a	n/a	n/a	n/a	0.0
	16-30		n/a	n/a	0.0	n/a	n/a	n/a	n/a	n/a	0.0
	31-45		n/a	n/a	0.1	n/a	n/a	n/a	n/a	n/a	0.1
	46-60		n/a	n/a	0.6	n/a	n/a	n/a	n/a	n/a	0.8
	61-75		n/a	n/a	4.1	n/a	n/a	n/a	n/a	n/a	6.4
	76+		n/a	n/a	5.4	n/a	n/a	n/a	n/a	n/a	23.6
Hungary	0-15	M	n/a	n/a	0.0	n/a	n/a	n/a	n/a	n/a	0.0
	16-30		n/a	n/a	0.0	n/a	n/a	n/a	n/a	n/a	0.0
	31-45		n/a	n/a	0.0	n/a	n/a	n/a	n/a	n/a	0.2
	46-60		n/a	n/a	0.9	n/a	n/a	n/a	n/a	n/a	2.3
	61-75		n/a	n/a	2.6	n/a	n/a	n/a	n/a	n/a	10.1
	76+		n/a	n/a	3.2	n/a	n/a	n/a	n/a	n/a	22.8
	0-15	F	n/a	n/a	0.0	n/a	n/a	n/a	n/a	n/a	0.0
	16-30		n/a	n/a	0.0	n/a	n/a	n/a	n/a	n/a	0.0
	31-45		n/a	n/a	0.1	n/a	n/a	n/a	n/a	n/a	0.1
	46-60		n/a	n/a	0.9	n/a	n/a	n/a	n/a	n/a	1.1
	61-75		n/a	n/a	2.0	n/a	n/a	n/a	n/a	n/a	5.3
	76+		n/a	n/a	1.4	n/a	n/a	n/a	n/a	n/a	12.6
Ireland	0-15	M	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	31-45		0.0	0.0	0.1	0.9	0.0	0.3	0.0	0.0	0.1
	46-60		0.0	0.3	1.1	4.1	0.0	4.2	0.1	0.1	1.1
	61-75		0.0	1.7	3.7	8.4	0.1	16.3	0.4	0.4	5.0
	76+		0.0	4.1	5.1	9.2	0.1	30.0	0.8	0.6	13.4

Continued on next page…

			Breast Cancer	Colo. Cancer	COPD	Diab.	Eso. Cancer	IHD	Lung Cancer	Oral Cancer	Stroke
	0-15	F	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	31-45		0.3	0.0	0.1	1.0	0.0	0.1	0.0	0.0	0.1
	46-60		2.0	0.3	1.1	3.2	0.0	2.0	0.1	0.0	0.9
	61-75		4.8	1.2	3.9	6.4	0.0	9.5	0.2	0.1	4.0
	76+		6.9	2.8	3.6	8.1	0.1	21.2	0.4	0.2	11.3
Italy	0-15	M	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	31-45		0.0	0.0	0.0	1.0	0.0	0.2	0.0	0.0	0.1
	46-60		0.0	0.4	1.2	4.2	0.0	2.5	0.1	0.2	1.0
	61-75		0.0	2.0	4.6	12.2	0.1	9.8	0.7	0.6	5.2
	76+		0.0	4.9	5.5	17.5	0.1	17.5	1.3	1.0	15.8
	0-15	F	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
	31-45		0.4	0.1	0.1	0.9	0.0	0.1	0.0	0.0	0.0
	46-60		2.6	0.4	0.9	3.6	0.0	1.2	0.1	0.1	0.5
	61-75		5.9	1.4	2.6	12.5	0.0	5.9	0.3	0.2	3.3
	76+		8.6	3.2	2.5	18.1	0.0	12.8	0.5	0.3	12.3
Latvia	0-15	M	0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0
	31-45		0.0	0.0	0.1	n/a	0.0	n/a	0.0	0.0	0.2
	46-60		0.0	0.1	0.6	n/a	0.0	n/a	0.1	0.1	2.5
	61-75		0.0	0.7	2.1	n/a	0.1	n/a	0.6	0.5	13.1
	76+		0.0	1.5	3.4	n/a	0.2	n/a	0.8	0.7	33.2
	0-15	F	0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0
	31-45		0.2	0.0	0.1	n/a	0.0	n/a	0.0	0.0	0.0
	46-60		1.0	0.1	0.6	n/a	0.0	n/a	0.0	0.0	0.0
	61-75		2.2	0.7	1.1	n/a	0.0	n/a	0.1	0.1	0.1
	76+		3.1	1.4	0.8	n/a	0.0	n/a	0.1	0.1	0.3
Lithuania	0-15	M	0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0
	31-45		0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.1
	46-60		0.0	0.2	0.8	n/a	0.0	n/a	0.2	0.1	1.5
	61-75		0.0	0.8	2.4	n/a	0.1	n/a	0.8	0.3	7.3
	76+		0.0	2.0	3.6	n/a	0.1	n/a	1.4	0.6	16.3
	0-15	F	0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0
	31-45		0.2	0.0	0.1	n/a	0.0	n/a	0.0	0.0	0.1
	46-60		1.0	0.2	0.6	n/a	0.0	n/a	0.0	0.0	1.0
	61-75		2.2	0.7	1.3	n/a	0.0	n/a	0.1	0.1	5.4

Continued on next page…

			Breast Cancer	Colo. Cancer	COPD	Diab.	Eso. Cancer	IHD	Lung Cancer	Oral Cancer	Stroke
	76+		3.1	1.3	1.1	n/a	0.0	n/a	0.2	0.2	12.9
Luxembourg	0-15	M	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	0.0
	16-30		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	0.0
	31-45		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	0.1
	46-60		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	1.1
	61-75		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	5.8
	76+		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	15.1
	0-15	F	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	0.0
	16-30		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	0.0
	31-45		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	0.1
	46-60		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	0.9
	61-75		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	4.6
	76+		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	13.0
Malta	0-15	M	0.0	0.0	n/a	n/a	0.0	n/a	0.0	0.0	0.0
	16-30		0.0	0.0	n/a	n/a	0.0	n/a	0.0	0.0	0.0
	31-45		0.0	0.1	n/a	n/a	0.0	n/a	0.0	0.0	0.1
	46-60		0.0	0.3	n/a	n/a	0.0	n/a	0.1	0.2	1.3
	61-75		0.0	0.9	n/a	n/a	0.0	n/a	0.4	0.5	6.4
	76+		0.0	1.9	n/a	n/a	0.1	n/a	0.6	1.0	16.2
	0-15	F	0.0	0.0	n/a	n/a	0.0	n/a	0.0	0.0	0.0
	16-30		0.0	0.0	n/a	n/a	0.0	n/a	0.0	0.0	0.0
	31-45		0.2	0.0	n/a	n/a	0.0	n/a	0.0	0.0	0.1
	46-60		1.7	0.2	n/a	n/a	0.0	n/a	0.0	0.1	0.7
	61-75		3.7	0.8	n/a	n/a	0.0	n/a	0.1	0.2	4.2
	76+		5.6	1.9	n/a	n/a	0.0	n/a	0.1	0.4	12.6
Netherlands	0-15	M	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	31-45		0.0	0.0	0.1	1.3	0.0	0.3	0.0	0.0	0.1
	46-60		0.0	0.3	1.5	5.7	0.0	3.5	0.1	0.2	0.9
	61-75		0.0	1.7	5.0	12.7	0.1	13.6	0.4	0.5	4.5
	76+		0.0	4.1	6.6	15.9	0.1	28.2	0.8	0.9	12.6
	0-15	F	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	31-45		0.5	0.0	0.1	1.0	0.0	0.1	0.0	0.0	0.1
	46-60		2.6	0.3	1.9	4.2	0.0	1.6	0.1	0.1	0.7
	61-75		5.7	1.3	4.8	11.7	0.0	7.8	0.2	0.3	3.5
	76+		8.6	3.0	5.3	17.4	0.0	18.9	0.2	0.4	10.1
Poland	0-15	M	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
	31-45		0.0	0.0	0.5	3.2	0.0	0.3	0.0	0.0	0.1

Continued on next page…

			Breast Cancer	Colo. Cancer	COPD	Diab.	Eso. Cancer	IHD	Lung Cancer	Oral Cancer	Stroke
	46-60		0.0	0.2	2.3	8.2	0.0	3.5	0.1	0.1	1.6
	61-75		0.0	0.8	8.8	12.5	0.0	14.3	0.4	0.4	7.6
	76+		0.0	1.7	8.8	12.6	0.0	25.2	0.5	0.7	15.0
	0-15	F	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
	31-45		0.2	0.0	0.7	3.2	0.0	0.1	0.0	0.0	0.1
	46-60		1.2	0.2	2.2	9.5	0.0	1.8	0.1	0.1	0.9
	61-75		3.1	0.7	4.9	16.1	0.0	9.4	0.2	0.1	4.9
	76+		4.4	1.1	4.9	16.7	0.0	19.2	0.4	0.2	11.7
Portugal	0-15	M	0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0
	31-45		0.0	0.1	0.0	n/a	0.0	n/a	0.0	0.1	0.3
	46-60		0.0	0.4	1.0	n/a	0.0	n/a	0.1	0.3	3.0
	61-75		0.0	1.8	3.5	n/a	0.1	n/a	0.3	0.6	14.3
	76+		0.0	3.6	3.7	n/a	0.1	n/a	0.4	1.0	37.8
	0-15	F	0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0
	31-45		0.4	0.1	0.0	n/a	0.0	n/a	0.0	0.0	0.1
	46-60		1.8	0.3	0.3	n/a	0.0	n/a	0.0	0.1	1.3
	61-75		3.6	1.1	1.1	n/a	0.0	n/a	0.1	0.2	8.1
	76+		5.0	2.2	1.0	n/a	0.0	n/a	0.1	0.3	27.0
Romania	0-15	M	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	16-30		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	31-45		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	46-60		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	61-75		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	76+		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	0-15	F	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	16-30		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	31-45		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	46-60		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	61-75		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	76+		n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Slovakia	0-15	M	0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0
	31-45		0.0	0.1	0.0	n/a	0.0	n/a	0.0	0.1	0.1
	46-60		0.0	0.4	0.9	n/a	0.0	n/a	0.2	0.4	1.0
	61-75		0.0	1.9	2.7	n/a	0.1	n/a	0.6	0.9	5.6
	76+		0.0	3.9	3.0	n/a	0.1	n/a	0.8	1.1	12.3
	0-15	F	0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	n/a	0.0	n/a	0.0	0.0	0.0

Continued on next page…

			Breast Cancer	Colo. Cancer	COPD	Diab.	Eso. Cancer	IHD	Lung Cancer	Oral Cancer	Stroke
	31-45		0.2	0.0	0.1	n/a	0.0	n/a	0.0	0.0	0.0
	46-60		1.2	0.2	0.8	n/a	0.0	n/a	0.0	0.0	0.5
	61-75		2.7	1.0	1.5	n/a	0.0	n/a	0.1	0.1	3.1
	76+		3.7	1.9	1.0	n/a	0.0	n/a	0.1	0.2	7.7
Slovenia	0-15	M	0.0	0.0	n/a	n/a	0.0	n/a	0.0	0.0	0.0
	16-30		0.0	0.0	n/a	n/a	0.0	n/a	0.0	0.0	0.0
	31-45		0.0	0.0	n/a	n/a	0.0	n/a	0.0	0.0	0.1
	46-60		0.0	0.3	n/a	n/a	0.0	n/a	0.2	0.3	1.1
	61-75		0.0	1.6	n/a	n/a	0.1	n/a	0.8	0.8	5.2
	76+		0.0	3.2	n/a	n/a	0.1	n/a	1.2	1.2	13.5
	0-15	F	0.0	0.0	n/a	n/a	0.0	n/a	0.0	0.0	0.0
	16-30		0.0	0.0	n/a	n/a	0.0	n/a	0.0	0.0	0.0
	31-45		0.2	0.0	n/a	n/a	0.0	n/a	0.0	0.0	0.1
	46-60		1.5	0.2	n/a	n/a	0.0	n/a	0.1	0.1	0.6
	61-75		3.3	1.0	n/a	n/a	0.0	n/a	0.2	0.2	3.5
	76+		4.8	1.9	n/a	n/a	0.0	n/a	0.3	0.3	10.6
Spain	0-15	M	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
	31-45		0.0	0.0	0.0	1.7	0.0	0.1	0.0	0.1	0.1
	46-60		0.0	0.3	1.2	7.5	0.0	1.9	0.2	0.4	1.0
	61-75		0.0	1.4	4.0	18.6	0.1	7.6	0.6	1.2	5.0
	76+		0.0	3.3	4.4	16.0	0.1	13.3	0.9	2.0	12.6
	0-15	F	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
	31-45		0.3	0.0	0.1	1.6	0.0	0.1	0.0	0.0	0.0
	46-60		1.7	0.2	1.1	4.9	0.0	0.9	0.0	0.1	0.5
	61-75		3.8	0.9	3.6	14.1	0.0	4.7	0.1	0.2	3.0
	76+		5.4	1.9	3.9	17.5	0.0	10.0	0.1	0.3	9.8
Sweden	0-15	M	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
	31-45		0.0	0.0	0.0	1.9	0.0	0.3	0.0	0.0	0.1
	46-60		0.0	0.2	1.2	6.0	0.0	3.9	0.1	0.1	1.0
	61-75		0.0	1.1	4.3	12.7	0.0	14.2	0.2	0.3	4.8
	76+		0.0	3.0	5.0	16.4	0.1	27.0	0.3	0.7	14.1
	0-15	F	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
	31-45		0.3	0.0	0.1	1.4	0.0	0.1	0.0	0.0	0.0
	46-60		2.1	0.2	1.3	3.8	0.0	1.9	0.1	0.1	0.6
	61-75		5.3	1.0	4.0	9.6	0.0	8.3	0.2	0.2	3.0
	76+		8.0	2.4	3.7	14.5	0.0	19.2	0.1	0.4	10.7

Continued on next page…

			Breast Cancer	Colo. Cancer	COPD	Diab.	Eso. Cancer	IHD	Lung Cancer	Oral Cancer	Stroke
UK	0-15	M	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	31-45		0.0	0.0	0.0	0.7	0.0	0.3	0.0	0.0	0.1
	46-60		0.0	0.2	0.8	4.0	0.0	3.9	0.1	0.1	1.1
	61-75		0.0	1.4	3.4	9.9	0.1	14.9	0.5	0.4	5.2
	76+		0.0	3.7	5.9	11.9	0.1	27.7	1.3	0.7	13.8
	0-15	F	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	16-30		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	31-45		0.3	0.0	0.1	0.8	0.0	0.1	0.0	0.0	0.1
	46-60		2.1	0.2	0.8	2.8	0.0	1.9	0.1	0.1	0.7
	61-75		5.0	1.0	2.8	6.9	0.0	8.6	0.3	0.2	3.9
	76+		7.6	2.6	3.7	9.1	0.1	19.5	0.7	0.3	12.4

C.2 Risk factor data

The DYNAMO-HIA project aimed to cover three major life-style-related risk factors: alcohol consumption, BMI, and smoking. The following Table C.3, shows country-specific data by sex aggregated into six age groups. For alcohol, the table shows the share of heavy drinkers, i.e. a daily consumption of 40 grams of pure alcohol or more. For BMI, the table shows the share of obese people, i.e. a BMI of 30 or more. For smoking, the table shows the share of smokers in the population.

Table C.3: Overview of the population-level prevalence of risk factors used in DYNAMO-HIA aggregated by age (in percent). [109,133,134]

		Alcohol (heavy drinkers)		BMI (obese)		Smoking (smokers)	
	Age	M	F	M	F	M	F
Austria	0-15	n/a	n/a	4.4	4.1	n/a	n/a
	16-30	n/a	n/a	13.1	6.6	n/a	n/a
	31-45	n/a	n/a	21.6	11.4	n/a	n/a
	46-60	n/a	n/a	24.4	23.0	n/a	n/a
	61-75	n/a	n/a	24.3	28.3	n/a	n/a
	76+	n/a	n/a	26.2	14.9	n/a	n/a
Belgium	0-15	n/a	n/a	4.4	4.1	0.0	0.0
	16-30	n/a	n/a	4.1	3.1	39.7	30.0
	31-45	n/a	n/a	8.7	7.4	42.3	31.9

Continued on next Page...

		Alcohol (heavy drinkers)		BMI (obese)		Smoking (smokers)	
	Age	M	F	M	F	M	F
	46-60	n/a	n/a	14.7	14.9	33.3	21.2
	61-75	n/a	n/a	17.2	17.3	20.8	8.4
	76+	n/a	n/a	18.5	13.1	11.5	3.5
Bulgaria	0-15	n/a	n/a	n/a	n/a	n/a	n/a
	16-30	n/a	n/a	n/a	n/a	n/a	n/a
	31-45	n/a	n/a	n/a	n/a	n/a	n/a
	46-60	n/a	n/a	n/a	n/a	n/a	n/a
	61-75	n/a	n/a	n/a	n/a	n/a	n/a
	76+	n/a	n/a	n/a	n/a	n/a	n/a
Cyprus	0-15	n/a	n/a	n/a	n/a	n/a	n/a
	16-30	n/a	n/a	n/a	n/a	n/a	n/a
	31-45	n/a	n/a	n/a	n/a	n/a	n/a
	46-60	n/a	n/a	n/a	n/a	n/a	n/a
	61-75	n/a	n/a	n/a	n/a	n/a	n/a
	76+	n/a	n/a	n/a	n/a	n/a	n/a
Czech Republic	0-15	0.1	0.2	3.1	2.5	0.0	0.0
	16-30	5.5	1.8	8.2	8.0	43.9	26.5
	31-45	11.5	1.2	18.4	15.9	43.3	34.4
	46-60	12.8	1.3	32.2	24.7	33.2	23.2
	61-75	9.7	1.0	32.6	39.2	18.9	9.8
	76+	6.3	0.7	27.7	21.4	14.5	1.2
Denmark	0-15	1.0	0.3	1.5	2.3	0.0	0.0
	16-30	14.5	3.4	4.5	4.1	38.3	34.4
	31-45	14.4	3.3	8.0	3.2	40.6	39.0
	46-60	17.3	4.7	12.1	8.8	42.5	36.5
	61-75	16.7	4.2	12.1	12.6	36.2	30.1
	76+	14.8	3.3	12.6	6.3	26.6	14.6
Estonia	0-15	0.4	0.0	n/a	n/a	0.0	0.0
	16-30	5.5	0.5	n/a	n/a	55.2	32.4
	31-45	6.7	0.6	n/a	n/a	58.0	30.4
	46-60	6.6	0.5	n/a	n/a	50.8	20.7
	61-75	3.7	0.2	n/a	n/a	40.4	11.0
	76+	1.0	0.0	n/a	n/a	32.3	8.5
Finland	0-15	0.3	0.1	3.3	3.9	0.0	0.0
	16-30	5.9	0.8	6.5	5.4	39.3	32.4
	31-45	6.9	0.9	14.0	11.5	37.5	25.4
	46-60	6.7	1.3	23.9	27.4	29.2	19.8

Continued on next Page…

		Alcohol (heavy drinkers)		BMI (obese)		Smoking (smokers)	
	Age	M	F	M	F	M	F
	61-75	8.1	0.6	33.0	35.6	24.1	12.8
	76+	9.7	0.1	34.2	19.4	22.4	11.5
France	0-15	0.1	0.0	3.3	3.6	0.0	0.0
	16-30	5.1	0.5	5.9	4.9	35.9	27.7
	31-45	11.6	0.7	9.6	9.4	32.2	25.5
	46-60	17.3	1.0	13.5	16.8	23.7	12.1
	61-75	13.5	1.5	15.3	18.3	10.5	4.5
	76+	8.2	1.2	15.8	11.6	7.2	1.2
Germany	0-15	0.9	0.2	4.4	4.1	0.0	0.0
	16-30	9.3	1.7	10.6	8.4	47.9	32.5
	31-45	7.8	1.3	15.6	16.4	37.8	26.8
	46-60	9.6	1.5	25.5	22.8	39.7	15.2
	61-75	10.6	1.6	30.2	32.5	29.8	7.8
	76+	8.8	1.3	23.2	20.1	25.2	6.3
Greece	0-15	1.0	0.0	n/a	n/a	n/a	n/a
	16-30	13.8	0.6	n/a	n/a	n/a	n/a
	31-45	14.4	0.5	n/a	n/a	n/a	n/a
	46-60	14.7	0.5	n/a	n/a	n/a	n/a
	61-75	11.3	0.3	n/a	n/a	n/a	n/a
	76+	6.5	0.2	n/a	n/a	n/a	n/a
Hungary	0-15	n/a	n/a	n/a	n/a	0.0	0.0
	16-30	n/a	n/a	n/a	n/a	47.7	37.3
	31-45	n/a	n/a	n/a	n/a	47.3	39.0
	46-60	n/a	n/a	n/a	n/a	34.6	21.4
	61-75	n/a	n/a	n/a	n/a	17.4	6.6
	76+	n/a	n/a	n/a	n/a	11.2	3.3
Ireland	0-15	0.7	0.1	6.3	7.2	0.0	0.0
	16-30	10.1	1.4	11.3	7.8	35.6	35.5
	31-45	9.2	1.1	20.8	13.3	36.7	29.2
	46-60	8.9	0.7	24.4	25.0	24.8	25.9
	61-75	5.7	0.3	24.0	33.2	25.1	22.8
	76+	1.3	0.3	25.6	17.8	25.3	9.4
Italy	0-15	0.0	0.0	6.1	6.4	0.0	0.0
	16-30	5.4	0.7	10.8	6.0	37.0	21.7
	31-45	11.9	1.4	12.5	9.7	38.3	25.1
	46-60	16.1	2.0	20.6	24.3	29.9	15.8
	61-75	15.1	1.7	19.4	29.0	18.4	6.3

Continued on next Page…

		Alcohol (heavy drinkers)		BMI (obese)		Smoking (smokers)	
	Age	M	F	M	F	M	F
	76+	11.9	1.3	22.1	17.2	9.3	2.0
Latvia	0-15	0.1	0.0	n/a	n/a	0.0	0.0
	16-30	3.0	0.4	n/a	n/a	58.9	35.4
	31-45	4.6	0.3	n/a	n/a	60.0	24.9
	46-60	3.8	0.2	n/a	n/a	46.5	10.3
	61-75	2.9	0.1	n/a	n/a	31.7	2.2
	76+	1.7	0.0	n/a	n/a	27.0	1.7
Lithuania	0-15	n/a	n/a	n/a	n/a	0.0	0.0
	16-30	n/a	n/a	n/a	n/a	50.7	27.3
	31-45	n/a	n/a	n/a	n/a	54.7	22.5
	46-60	n/a	n/a	n/a	n/a	39.8	9.7
	61-75	n/a	n/a	n/a	n/a	31.9	4.1
	76+	n/a	n/a	n/a	n/a	27.1	3.3
Luxem bourg	0-15	n/a	n/a	n/a	n/a	n/a	n/a
	16-30	n/a	n/a	n/a	n/a	n/a	n/a
	31-45	n/a	n/a	n/a	n/a	n/a	n/a
	46-60	n/a	n/a	n/a	n/a	n/a	n/a
	61-75	n/a	n/a	n/a	n/a	n/a	n/a
	76+	n/a	n/a	n/a	n/a	n/a	n/a
Malta	0-15	1.5	0.6	n/a	n/a	n/a	n/a
	16-30	16.3	4.5	n/a	n/a	n/a	n/a
	31-45	13.1	1.5	n/a	n/a	n/a	n/a
	46-60	11.5	0.8	n/a	n/a	n/a	n/a
	61-75	8.3	0.6	n/a	n/a	n/a	n/a
	76+	4.3	0.7	n/a	n/a	n/a	n/a
Nether-lands	0-15	0.6	0.1	2.2	2.8	0.0	0.0
	16-30	9.0	1.1	4.7	5.2	39.1	32.6
	31-45	10.2	1.8	8.5	9.6	40.1	32.1
	46-60	11.5	3.0	15.1	14.4	33.8	25.7
	61-75	8.9	2.3	10.2	19.0	25.0	17.0
	76+	4.3	1.7	10.4	10.4	19.2	8.4
Poland	0-15	0.4	0.0	5.0	4.8	n/a	n/a
	16-30	4.3	0.1	4.2	3.4	n/a	n/a
	31-45	4.1	0.2	15.0	13.9	n/a	n/a
	46-60	3.8	0.2	23.9	25.6	n/a	n/a
	61-75	2.6	0.1	18.8	40.4	n/a	n/a

Continued on next Page…

		Alcohol (heavy drinkers)		BMI (obese)		Smoking (smokers)	
	Age	M	F	M	F	M	F
	76+	1.2	0.0	21.0	23.2	n/a	n/a
Portugal	0-15	n/a	n/a	8.7	8.4	0.0	0.0
	16-30	n/a	n/a	9.0	5.4	44.8	19.4
	31-45	n/a	n/a	17.2	15.1	44.2	13.9
	46-60	n/a	n/a	20.4	22.3	28.3	4.3
	61-75	n/a	n/a	19.3	27.7	14.4	1.0
	76+	n/a	n/a	23.2	23.5	6.8	0.7
Romania	0-15	n/a	n/a	n/a	n/a	n/a	n/a
	16-30	n/a	n/a	n/a	n/a	n/a	n/a
	31-45	n/a	n/a	n/a	n/a	n/a	n/a
	46-60	n/a	n/a	n/a	n/a	n/a	n/a
	61-75	n/a	n/a	n/a	n/a	n/a	n/a
	76+	n/a	n/a	n/a	n/a	n/a	n/a
Slovakia	0-15	n/a	n/a	n/a	n/a	0.0	0.0
	16-30	n/a	n/a	n/a	n/a	44.3	31.7
	31-45	n/a	n/a	n/a	n/a	45.8	30.6
	46-60	n/a	n/a	n/a	n/a	30.7	13.3
	61-75	n/a	n/a	n/a	n/a	16.4	3.0
	76+	n/a	n/a	n/a	n/a	13.8	2.2
Slovenia	0-15	n/a	n/a	n/a	n/a	n/a	n/a
	16-30	n/a	n/a	n/a	n/a	n/a	n/a
	31-45	n/a	n/a	n/a	n/a	n/a	n/a
	46-60	n/a	n/a	n/a	n/a	n/a	n/a
	61-75	n/a	n/a	n/a	n/a	n/a	n/a
	76+	n/a	n/a	n/a	n/a	n/a	n/a
Spain	0-15	0.1	0.1	9.0	7.6	0.0	0.0
	16-30	4.2	1.0	6.7	3.3	40.9	29.0
	31-45	8.9	1.1	10.7	10.8	41.5	31.4
	46-60	12.4	1.4	18.3	29.2	32.2	21.8
	61-75	10.8	1.0	25.1	37.0	13.9	14.7
	76+	7.7	0.1	31.0	40.0	5.4	20.6
Sweden	0-15	0.7	0.2	3.9	3.6	0.0	0.0
	16-30	9.9	2.0	5.2	6.4	14.2	19.7
	31-45	9.3	2.0	13.4	10.3	19.2	23.5
	46-60	6.6	1.7	13.2	10.9	21.5	24.8
	61-75	4.4	1.0	22.0	16.6	13.4	13.1

Continued on next Page…

		Alcohol (heavy drinkers)		BMI (obese)		Smoking (smokers)	
	Age	M	F	M	F	M	F
	76+	3.7	0.5	22.8	8.5	4.7	5.2
UK	0-15	1.6	1.1	6.4	8.9	0.0	0.0
	16-30	26.4	13.9	11.4	14.8	34.7	32.4
	31-45	31.0	13.0	24.8	24.3	28.7	27.9
	46-60	31.5	13.1	28.6	28.8	22.3	22.1
	61-75	24.0	8.4	27.9	30.9	15.4	15.7
	76+	16.3	5.3	16.4	21.3	7.0	6.2

C.3 Relative risks

The following tables report the relative risk that connect the three risk factors causally to disease incidence and total mortality. In addition, the relative risk for stroke and IHD for diabetes patients are shown.

In addition, the DYNAMO-HIA model can use relative risks to the target disease by time since stopping (by age and sex). To derive the necessary values, we implemented the approach used by Hoogenveen et al.[51] 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:

Table C.4: Overview of relative risks from diabetes to IHD and stroke in DYNAMO-HIA.[135–137]

	Males	Females
Diabetes to IHD		
Persons aged up to 55	2.66	3.53
Persons aged 56+	1.93	2.59
Diabetes to stroke		
Persons aged up to 49	2	2.9
Persons aged 50+	1.8	2.2

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

where

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

The values for γ_0 and θ are taken from Hoogenveen et al.[51]

Table C.5: Relative risks of smoking on mortality and diseases used in DYNAMO-HIA. [138–141]

Outcome	Males aged 35 and above			Females aged 35 and above		
	Never	Current	Former	Never	Current	Former
All-cause mortality						
Persons aged 35+	1	2.07	1.35	1	1.74	1.23
Oral Cancer						
Persons aged 35+	1	10.89	3.4	1	5.08	2.29
Esophagus Cancer						
Persons aged 35+	1	6.76	4.46	1	7.75	2.79
Lung cancer						
Persons aged 35-39	1	1.3	1	1	2	1
Persons aged 40-44	1	1	1	1	1	1
Persons aged 45-49	1	5.78	2.37	1	18.08	8.07
Persons aged 50-54	1	24.97	10.7	1	11.14	3.28
Persons aged 55-59	1	34.02	11.66	1	17.87	5.33
Persons aged 60-64	1	31.47	11.71	1	13.32	4.91
Persons aged 65+	1	28.4	9.7	1	17.49	5.54
IHD						
Persons aged 35-39	1	3.25	1.21	1	1	1.44
Persons aged 40-44	1	4.71	1.15	1	1.89	2.25
Persons aged 45-49	1	5.85	2.03	1	7.71	2.08
Persons aged 50-54	1	3.69	1.93	1	5.69	2.95
Persons aged 55-59	1	2.71	1.64	1	3.06	1.19
Persons aged 60-64	1	2.39	1.58	1	2.56	1.08
Persons aged 65+	1	1.91	1.4	1	2.48	1.22
Stroke						
Persons aged 35-39	1	1	1	1	2	1
Persons aged 40-44	1	1.05	1	1	5.67	2.25
Persons aged 45-49	1	3.75	1	1	8.22	1.19
Persons aged 50-54	1	6.08	2.24	1	4.58	1.38
Persons aged 55-59	1	3.96	1.14	1	5.77	1.22
Persons aged 60-64	1	2.55	1.01	1	2.76	1.28
Persons aged 65+	1	2.69	1.29	1	2.58	1.14
COPD						
Persons aged 35-49	1	1	1	1	1	1
Persons aged 50-54	1	8.13	3.06	1	12.92	7.39
Persons aged 55-59	1	9.8	8.25	1	9.47	5.55
Persons aged 60-64	1	13.21	12.65	1	11.19	6.63
Persons aged 65+	1	18.93	11.92	1	14.72	9.73

Table C.6: Relative risks of BMI (categorical) on mortality and diseases used in DYNAMO-HIA. [116,142–170]

Outcome (x = multiplier of differential risk for age adjustment)	RR overweight BMI 25-29.9		RR obesity BMI 30 or more	
	Normal weight = 1		Normal weight = 1	
	males	females	males	females
All-cause mortality	1.2	1.15	1.55	1.5
x 0.98 from age 50				
x 0.95 from age 60				
x 0.90 from age 70				
IHD	1.35	1.35	2	2
x 0.70 from age 65				
Stroke	1.2	1.2	1.5	1.55
x 0.75 from age 65				
Diabetes	2.25	2.3	5.5	7
x 0.92 from age 60				
x 0.90 from age 75				
COPD	1	1	1	1
Lung cancer	0.8	0.88	0.65	0.7
Breast cancer	n/a	1	n/a	1
over age 50		1.12		
Oral cancer	0.8	0.88	0.65	0.7
Colorectal cancer	1.2	1.08	1.4	1.1
x 0.90 from age 45				
Esophagal cancer	1	1	1	1

131,06 mm

Appendix C

Table C.7: Relative risks of BMI (continuous) on mortality and diseases used in DYNAMO-HIA. [116,142–170]

Outcome	RR per unit BMI above BMI 22	
(x = multiplier of differential risk for age adjustment)	males	females
All-cause mortality	1.07	1.03
x 0.98 from age 50 x 0.95 from age 60 x 0.90 from age 70		
IHD	1.07	1.1
x 0.70 from age 65		
Stroke	1.04	1.04
x 0.75 from age 65		
Diabetes	1.18	1.22
x 0.92 from age 60 x 0.90 from age 75		
COPD	1	1
Lung cancer	0.97	0.98
Breast cancer	n/a	1
x 1.02 from age 50		
Oral cancer	0.97	0.98
Colorectal cancer	1.04	1.02
x 0.90 over age 45		
Esophagal cancer (all forms combined)	1	1

Table C.8: Relative risks of alcohol consumption on mortality and diseases used in DYNAMO-HIA.[171–176]

Outcome	Males aged 15 years and over Drinking categories (grams per day)					Females aged 15 years and over Drinking categories (grams per day)				
	0 - <0.25	0.25 - <20	20 - <40	40 - <60	<60	0 - <0.25	0.25 - <20	20 - <40	40 - <60	<60
All-cause mortality										
Persons aged 16-24	1	1.07	1.25	1.48	1.88	1	1.04	1.17	1.31	1.58
Persons aged 25-34	1	1.05	1.21	1.40	1.75	1	1.04	1.15	1.29	1.54
Persons aged 35-44	1	1	1.10	1.23	1.47	1	1.03	1.15	1.30	1.56
Persons aged 45-54	1	0.96	1.01	1.1	1.26	1	1.02	1.13	1.26	1.51
Persons aged 55-64	1	0.94	0.98	1.04	1.16	1	1	1.09	1.22	1.46
Persons aged 65-74	1	0.94	0.97	1.02	1.11	1	0.99	1.06	1.17	1.38
Persons aged 75-84	1	0.95	0.97	1.02	1.11	1	0.98	1.05	1.15	1.35
Persons aged 85-95	1	0.96	0.98	1.02	1.09	1	0.98	1.03	1.12	1.27
IHD	1	0.82	0.82	0.87	1.13	1	0.82	0.82	0.87	1.13
Stroke	1	0.91	1.01	1.18	1.55	1	0.70	0.79	1.08	2.74
Diabetes mellitus	1	0.72	0.86	1	1	1	0.72	0.86	1	1
COPD	1	1	1	1	1	1	1	1	1	1
Lung cancer	1	1	1	1	1	1	1	1	1	1
Colorectal cancer	1	1	1.08	1.30	1.72	1	1	1.11	1.33	1.62
Oral cancer	1	1.31	2.08	3.02	4.32	1	1.33	2.18	3.26	4.85
Breast cancer	1	1	1	1	1	1	1	1.23	1.42	1.68
Esophageal cancer	1	1.17	1.61	2.19	3.18	1	1.17	1.61	2.19	3.18

D

PhD-Portfolio

Obtained academic degrees

- *Diplom-Politologe* (Freie Universität zu Berlin, 2003)
- Bachelor of Science in Statistics (Humboldt Universität Berlin, 2004)
- Master in Public Policy (Duke University, 2006)
- Postgraduate Diploma in Social Science Data Analysis (University of Essex, 2007)
- European Research Master in Demography (Max-Planck-Institute for Demographic Research, 2007)
- Master of Science in Statistics (Humboldt Universität zu Berlin, 2008)

General courses

- Biomedical English Writing and Communication (Erasmus University, 2009, 4 ECTS)

Specific courses

- Health Technology Assessment (Erasmus University, 2009, 5 ECTS)
- Advanced Economic Evaluation (Erasmus University, 2009, 5 ECTS)
- CVD Epidemiology and Modeling (Marie Curie course, 2008, 37.5hrs)
- Major Determinants and Major Diseases (NIHES, 2008, 1.4 ECTS)
- Multi-state models and models for competing risks (Erasmus MC, 2010, 0.6 ECTS)
- Craft of Smoothing (Erasmus MC, 2010, 0.6 ECTS)
- Health Issues in Humanitarian Crises (University Bielefeld, 2010, 1.5 ECTS)

Seminars and workshops

- Invited presentation at Quantitative HIA workshop at EUPHA 09 (Lodz, 2009, 40 hours)
- DYNAMO-HIA consortium meeting (Rotterdam, 2008, 40 hours)
- DYNAMO-HIA workshop at HIA 09 (Rotterdam, 2009, 40 hours)
- DYNAMO-HIA workshop at EUPHA 10 (Amsterdam, 2010, 40 hours)

Other

- Convenor of MMDV-Meetings at Erasmus MC 2008-2010 (60 hours)

Presentations

- "Spatial Modeling of all-cause mortality", Methodology Seminar (Erasmus MC, 2008, 20 hours)
- "Pitfalls of PIF", Methodology Seminar (Erasmus MC, 2009, 20 hours)
- MGZ Research Seminar (Erasmus MC, 2011, 20 hours)

(Inter)national conferences

- Paper presented at HIA 08 conference (Liverpool, 2008, 40 hours)
- Paper presented at German Statistical Week (Cologne, 2008, 40 hours)
- Key note lecture (co-authored) at HIA 10 (Rotterdam, 2010, 40 hours)
- Paper presented at EUPHA 09 (Lodz, 2009, 40 hours)
- Poster presented at EUPHA 10 (Amsterdam, 2010, 40 hours)
- Paper presented at DGEpi (Berlin, 2010, 40 hours)

Lecturing

- "Global Burden of Disease study" (Lecture and in-class exercise), International Institute of Social Studies (Den Haag, 2011, 40 hours)
- Medical Demography (Lecture), NIHES (Rotterdam, 2010, 20 hours)
- Medical Demography (Lecture), NIHES (Rotterdam, 2011, 40 hours)

Supervising practicals and excursions, Tutoring

- "Global Burden of Disease study" (in-class exercise), International Institute of Social Studies (Den Haag, 2010, 20 hours)

E

CV

Stefan K. Lhachimi was born on November 19th, 1976 in Bernburg (Germany). After obtaining his *Abitur* at the *Privates Don Bosco Gymnasium* in Essen under the auspices of the order of the Salesian Society, he went on to study political science in Berlin. From 2000 to 2004 he was employed as a research assistant at the Institute for Social Science, Humboldt University, working in the research unit for public policy and public administration (Prof. H. Wollmann) and, subsequently, at the research unit for the comparative analysis of political systems (Prof. Immergut). In 2003, he graduated with a *Diplom* in political science from *Freie Universität Berlin* and went on to obtain an intermediate degree in statistics from Humboldt University in 2004.

From 2004 to 2006, Stefan Lhachimi was a Max-Planck-Fellow at the Terry-Sanford-School of Public Policy, Duke University. He graduated Master of Public Policy with a thesis on the health effects of the liberalization of the Swedish alcohol market. In 2006/7, he received advanced training in demography at the Max-Planck-Institute for Demographic Research and started in 2007 to work at the Department of Public Health, Erasmus MC. In 2008, he obtained a Master of Science in statistics at Humboldt University and became later that year a fellow of the Royal Statistical Society. Furthermore, he collected extensive experience as a teaching fellow/instructor, inter alia, at Humboldt University, University of Essex, Duke University, Institute for Social Science (The Hague), and NIHES/Erasmus MC. Currently, Stefan Lhachimi works as a senior researcher at the Institute for Medical Technology Assessment, Erasmus University Rotterdam.

F

References

[1] European Centre for Health Policy. Health impact assessment: main concepts and suggested approach. Gothenburg consensus paper, 1999.

[2] J. Lennert Veerman. *Quantitative Health Impact Assessment: An exploration of methods and validity: PhD Thesis*. EUR, Rotterdam, 2007.

[3] J. M. Parry, J. R. Kemm, and on behalf of all participants of the Evaluation of Health Impact Assessment Workshop. Criteria for use in the evaluation of health impact assessments. *Public Health*, 119(12):1122–1129, 2005.

[4] J. Kemm. What is HIA and why might it be useful? In Matthias Wismar, Julia Blau, Kelly Ernst, and Josep Figueras, editors, *The Effectiveness of Health Impact Assessment. Scope and limitations of supporting decision-making in Europe*, European Observatory on Health Systems and Policies, pages 3–13. WHO, 2007.

[5] J. M. Parry and J. Wright. Community participation in health impact assessments: intuitively appealing but practically difficult. *Bullentin of World Health Organization*, 81(6):388, 2003.

[6] Karen Lock and Martin McKee. Health impact assessment: assessing opportunities and barriers to intersectoral health improvement in an expanded European Union. *J Epidemiol Community Health*, 59(5):356–360, 2005.

[7] Project Group. European Policy Health Impact Assessment, A Guide, 2004.

[8] J. Mindell, E. Ison, and M. Joffe. A glossary for health impact assessment. *J Epidemiol Community Health*, 57(9):647–651, 2003.

[9] B. L. Cole and J. E. Fielding. Health impact assessment: A tool to help policy makers understand health beyond health care. *Annual Review of Public Health*, 28:393–412, 2007.

[10] J. P. Mackenbach and Stefan K Lhachimi. Quantitative Health Impact Assessment: Where do we go from here?, 15th October, 2009.

[11] J. S. Mindell, A. Boltong, and I. Forde. A review of health impact assessment frameworks. *Public Health*, 122(11):1177–1187, 2008.

[12] Brian L. Cole, Riti Shimkhada, Jonathan E. Fielding, Gerald Kominski, and Hal Morgenstern. Methodologies for realizing the potential of health impact assessment. *American Journal of Preventive Medicine*, 28(4):382–389, 2005.

[13] Hal Morgenstern and Elaine S. Bursic. A method for using epidemiologic data to estimate the potential impact of an intervention on the health status of a target population. *Journal of Community Health*, 7(4):292–309, 1982.

[14] Kenneth J. Rothman, Sander Greenland, and Timothy L. Lash. *Modern epidemiology*. Lippincott Williams & Wilkins, Philadelphia, 2008.

[15] Debbie Abrahams, Fiona Haigh, and Andrew Pennington. A health impact assess-

ment of the European employment strategy across the European Union, August 2004.

[16] Debbie Abrahams, Lea den Broeder, Cathal Doyle, Rainer Fehr, Fiona Haigh, Odile Mekel, O. Metcalfe, Andrew Pennington, and A. Scott-Samuel. Policy Health Impact Assessment for the European Union: Final Project Report, August 2004.

[17] Fiona Haigh and Odile Mekel. Policy Health Impact Assessment for the European Union: Pilot Health Impact Assessment of the European Employment Strategy in Germany, August 2004.

[18] Rainer Fehr, Odile Mekel, Martin Lacombe, and Ulrike Wolf. Towards health impact assessment of drinking-water privatization — the example of waterborne carcinogens in North Rhine-Westphalia. *Bull World Health Organ*, 81(6):408–414, 2003.

[19] Mark McCarthy and Martin Utley. Quantitative Approaches to Health Impact Assessment. In John Kemm, Jayne Parry, and Stephen Palmer, editors, *Health Impact Assessment: Concepts, Theory, Techniques and Applications*. OUP, Oxford, 2004.

[20] Johanna M. Mooy and Lousie J. Gunning-Schepers. Computer-assisted health impact assessment for intersectoral health policy. *Health Policy*, 57:169–177, 2001.

[21] Michael C. Wolfson. POHEM - A Framework for Understanding and Modelling the Health of Human Populations. *World Health Stat Q*, 47(3-4):157–176, 1994.

[22] Martin Utley, Steve Gallivan, Jane Biddulph, Mark McCarthy, and Jake Ferguson. ARMADA – A Computer Model of the Impact of Environmental Factors on Health. *Health Care Manag Sci*, 6(3):137–146, 2003.

[23] Brian L Cole, Riti Shimkhada, Hal Morgenstern, Gerald Kominski, Jonathan E Fielding, and Sheng Wu. Projected health impact of the Los Angeles City living wage ordinance. *J Epidemiol Community Health*, 59(8):645–650, 2005.

[24] Klim McPherson, Tim Marsh, and J. Brown. Tackeling obesities: Future choices - Modelling future trends in obesity and the impact on health: Foresight Obesity Project, 2007.

[25] Matthias Wismar, Julia Blau, Kelly Ernst, Eva Elliott, Alison Golby, Loes van Herten, Teresa Lavin, Marius Sticka, and Gareth Williams. Implementation and institutionalizing HIA in Europe. In Matthias Wismar, Julia Blau, Kelly Ernst, and Josep Figueras, editors, *The Effectiveness of Health Impact Assessment. Scope and limitations of supporting decision-making in Europe*, European Observatory on Health Systems and Policies, pages 57–78. WHO, 2007.

[26] Rebecca Salay and Paul Lincoln. Health impact assessments in the European Union. *Lancet*, 372:860–861, 2008.

[27] J. D. F. Habbema, R. Boer, and J. J. Barendregt. Chronic Disease Modeling. In Kris Heggenhougen and Stella Quah, editors, *International Encyclopedia of Public Health*, volume 1, pages 704–709. Academic Press, San Diego, 2008.

[28] J. Kemm. The future challenges for HIA. *Environ Impact Assess Rev*, 25(7-8):799–807, 2005.

[29] J. L. Veerman, J. J. Barendregt, and J. P. Mackenbach. Quantitative health impact assessment: current practice and future directions. *J Epidemiol Community Health*, 59(5):361–370, 2005.

[30] M. Joffe and J. Mindell. A framework for the evidence base to support Health Impact Assessment. *J Epidemiol Community Health*, 56(2):132–138, 2002.

[31] J. Mindell and A. Boltong. Public Health Observatories: Supporting health impact assessment in practice. *Public Health*, 119(4):246–252, 2005.

[32] J. R. Kemm. Can Health Impact Assessment fulfil the expectations it raises? *Public Health*, 114(6):431–433, 2000.

[33] Matthias Wismar, Julia Blau, Kelly Ernst, and Josep Figueras, editors. *The Effectiveness of Health Impact Assessment. Scope and limitations of supporting decision-making in Europe*. European Observatory on Health Systems and Policies. WHO, 2007.

[34] Andrew L. Dannenberg, Rajiv Bhatia, Brian L. Cole, Sarah K. Heaton, Jason D. Feldman, and Candace D. Rutt. Use of Health Impact Assessment in the U.S.: 27 Case Studies, 1999-2007. *Am J Prev Med*, 34(3):241–256, 2008.

[35] Olaf Gefeller. An Annotated Bibliography on the Attributable Risk. *Biom J*, 34(8):1007–1012, 1992.

[36] H Brønnum-Hansen. How good is the Prevent model for estimating the health benefits of prevention? *J Epidemiol Community Health*, 53(5):300–305, 1999.

[37] Jennifer Mindell, Anna Hansell, David Morrison, Margaret Douglas, and Michael Joffe. What do we need for robust, quantitative health impact assessment? *J Public Health*, 23(3):173–178, 2001.

[38] Bjarne Robberstad. QALYs vs DALYs vs LYs gained: What are the differences, and what difference do they make for health care priority setting? *Norsk Epidemiologi*, 15(2):183–191, 2005.

[39] Ari Rabl. Analysis of air pollution mortality in term of life expectancy changes: relation between time series, intervention, and cohort studies. *Eviron Health*, 5(1), 2006.

[40] Frans Willekens. Biographic Forecasting: Bridging the Micro-Macro Gap in Population Forecasting. *N Z Popul Rev*, 31(1):77–124, 2005.

[41] Jonathan Karnon. Alternative decision modelling techniques for the evaluation of health care technologies: Markov processes versus discrete event simulation. *Health Econ*, 12(10):837–848, 2003.

[42] Alan D. Lopez, C. D. Mathers, M. Ezzati, Dean T. Jamison, and Christopher J. L. Murray. *Global burden of disease and risk factors*. Oxford University Press and World Bank, New York, NY Washington, DC, 2006.

[43] Majid Ezzati, Alan D. Lopez, Anthony Rodgers, Stephen Vander Hoorn, and Christopher J. L. Murray. Selected major risk factors and global and regional burden of disease. *Lancet*, 360(9343):1347–1360, 2002.

[44] Christopher J. L. Murray and Alan D. Lopez. On the Comparable Quantification of Health Risks: Lessons from the Global Burden of Disease Study. *Epidemiology*, 10(5):594–605, 1999.

[45] B.P. Will, J.-M. Berthelot, K.M. Nobrega, W. Flanagan, and W.K. Evans. Canada's Population Health Model (POHEM): a tool for performing economic evaluations of cancer control interventions. *Eur J Cancer*, 37(14):1797–1804, 2001.

[46] Bhash Naidoo, Margaret Thorogood, Klim McPherson, and Lousie J. Gunning-Schepers. Modelling the effects of increased physical activity on coronary heart disease in England and Wales. *J Epidemiol Community Health*, 51(2):144–150, 1997.

[47] J. J. Barendregt, G. J. van Oortmarssen, van Hout Ben A., Jacqueline M. van den Bosch, and Luc Bonneux. Coping with multiple morbidity in a life table. *Math Popul Stud*, 7(1):29–49, 1998.

[48] J. L. Veerman, J. J. Barendregt, E. F. van Beeck, J C Seidell, and J. P. Mackenbach. Stemming the obesity epidemic: a tantalizing prospect. *Obesity*, 15(9):2365–2370, 2007.

[49] J. L. Veerman, J. J. Barendregt, J. P. Mackenbach, and J Brug. Using epidemiological models to estimate the health effects of diet behaviour change: the example of tailored fruit and vegetable promotion. *Public Health Nutr*, 9(4):415–420, 2006.

[50] Rudolf Hoogenveen, A. E. de Hollander, and M. L. van Genugten. The chronic diseases modelling approach, 1998.

[51] Rudolf T Hoogenveen, Pieter H van Baal, Hendriek C Boshuizen, and Talitha L Feenstra. Dynamic effects of smoking cessation on disease incidence, mortality and quality of life: The role of time since cessation. *Cost Eff Resour Alloc*, 6(1):1, 2008.

[52] Pieter H. M. van Baal, Johan J. Polder, G. Ardine de Wit, Rudolf T. Hoogenveen, Talitha L. Feenstra, Hendriek C. Boshuizen, Peter M. Engelfriet, and Werner B. F. Brouwer. Lifetime Medical Costs of Obesity: Prevention No Cure for Increasing Health Expenditure. *PLoS Medicine*, 5(2):e29 EP, 2008.

[53] Wilbert O. Galitz. *The essential guide to user interface design: An introduction to GUI design principles and techniques*. Wiley, Indianapolis, Ind., 2007.

[54] Jan J. Barendregt. PREVENT V. 3.0: Work in Progress, nd.

[55] Faline Howes, Jodie Doyle, Nicki Jackson, and Elizabeth Waters. Evidence-based public health: the importance of finding 'difficult

to locate' public health and health promotion intervention studies for systematic reviews. *J Public Health*, 26(1):101–104, 2004.

[56] Mark Sculpher, Elisabeth Fenwick, and Karl Claxton. Assessing Quality in Decision Analytic Cost-Effectivness Models: A Suggested Framework and Example of Applications. *Pharmacoeconomics*, 17(5):461–477, 2000.

[57] Zoe Philips, Laura Bojke, Mark Sculpher, Karl Claxton, and Su Golder. Good Practice Guidelines for Decision-Analytic Modelling in Health Technology Assessment: A Review and Consolidation of Quality Assessment. *Pharmacoeconomics*, 24(4):355–371, 2006.

[58] Stefan K Lhachimi, Wilma J Nusselder, Pieter v. Baal, Hendriek C Boshuizen, Sido Mylius, and Jet Smit. DYNAMO-HIA Deliverable D1 - Model specification for a dynamic model for health impact assessment, 15th July 2008.

[59] M. McCarthy, J. P. Biddulph, M. Utley, J. Ferguson, and S. Gallivan. A health impact assessment model for environmental changes attributable to development projects. – *J Epidemiol Community Health*, 56(8):611–616, 2002.

[60] Pieter van Baal, Rudolf Hoogenveen, G. Ardine de Wit, and Hendriek Boshuizen. Estimating health-adjusted life expectancy conditional on risk factors: results for smoking and obesity, 2006.

[61] Linda Kok, Peter Engelfriet, Monique A. M. Jacobs-van der Bruggen, Rudolf T Hoogenveen, Hendriek C Boshuizen, and Monique W. M. Verschuren. The cost-effectiveness of implementing a new guideline for cardiovascular risk management in primary care in the Netherlands. *European journal of cardiovascular prevention and rehabilitation : official journal of the European Society of Cardiology, Working Groups on Epidemiology & Prevention and Cardiac Rehabilitation and Exercise Physiology*, 16(3):371–376, 2009.

[62] Stefan K Lhachimi, Wilma J Nusselder, Hendriek C Boshuizen, and J. P. Mackenbach. No Standard Tool for Quantification in Health Impact Assessment: A review. *Am J Prev Med*, 38(1):78–84, 2010.

[63] Dana P. Goldman, Baoping Shang, Jayanta Bhattacharya, Alan M. Garber, Michael Hurd, Geoffrey F. Joyce, Darius N. Lakdawalla, Constantijn Panis, and Paul G. Shekelle. Consequences of health trends and medical innovation for the future elderly. *Health affairs (Project Hope)*, 24 Suppl 2:W5R5–17, 2005.

[64] Jeroen van Meijgaard, Jonathan E Fielding, and Gerald F. Kominski. Assessing and forecasting population health: integrating knowledge and beliefs in a comprehensive framework. *Public health reports (Washington, D.C. : 1974)*, 124(6):778–789, 2009.

[65] Earl S. Ford, Umed A. Ajani, Janet B. Croft, Julia A. Critchley, Darwin R. Labarthe, Thomas E. Kottke, Wayne H. Giles, and Simon Capwell. Explaining the decrease in U.S-deaths from coronary heart disease, 1980-2000. – *The New England Journal of Medicine*, 356:2388–2398, 2007.

[66] Rudolf T. Hoogenveen, Pieter H. M. van Baal, and Hendriek C. Boshuizen. Chronic disease projections in heterogeneous ageing populations: approximating multi-state models of joint distributions by modelling marginal distributions. *Math Med Biol*, 2009.

[67] Juha Alho and Bruce D. Spencer. *Statistical demography and forecasting*. Springer series in statistics. Springer, New York ; London, 2005.

[68] Steven M. Shechter, Andrew J. Schaefer, R. Scott Braithwaite, and Mark S. Roberts. Increasing the efficiency of Monte Carlo cohort simulations with variance reduction techniques. *Medical decision making : an international journal of the Society for Medical Decision Making*, 26(5):550–553, 2006.

[69] Jan Barendregt, Gerrit van Oortmarssen, Theo Vos, and Christopher Murray. A generic model for the assessment of disease epidemiology: the computational basis of DisMod II. *Population Health Metrics*, 1(1):4, 2003.

[70] L. Andronis, P. Barton, and S. Bryan. Sensitivity analysis in economic evaluation: an audit of NICE current practice and a review of its use and value in decision-making. *Health technology assessment (Winchester, England)*, 13(29):iii, ix–xi, 1–61, 2009.

[71] A. Briggs, M. Sculpher, and M. Buxton. Uncertainty in the economic evaluation of health care technologies: the role of sensitivity analysis. *Health economics*, 3(2):95–104, 1994.

[72] Hendriek C Boshuizen and Pieter H. M. van Baal. Probabilistic sensitivity analysis: be a Bayesian. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*, 12(8):1210–1214, 2009.

[73] Tuhina Neogi and Yuqing Zhang. Re: "Easy SAS calculations for risk or prevalence ratios and differences". *American journal of epidemiology*, 163(12):1157; author reply 1159–61, 2006.

[74] J. M. Alho. On prevalence, incidence, and duration in general stable populations. *Biometrics*, 48(2):587–592, 1992.

[75] Jan vd Kassteele, Rudolf T. Hoogenveen, Pieter van Baal, and Hendriek C Boshuizen. Estimating net transition probabilities from cross-sectional data with application to risk factors in chronic disease modeling. manuscript. manuscript, 2011.

[76] Jeremy Lauer, Klaus Rohrich, Harald Wirth, Claude Charette, Steve Gribble, and Christopher Murray. PopMod: a longitudinal population model with two interacting disease states. *Cost effectiveness and resource allocation*, 1(1):6, 2003.

[77] P.M. Clarke, A.M. Gray, A. Briggs, A.J. Farmer, P. Fenn, R.J. Stevens, D.R. Matthews, I.M. Stratton, R.R. Holman, and on behalf of the UK Prospective Diabetes StudyGroup. A model to estimate the lifetime health outcomes of patients with Type 2 diabetes: the United Kingdom Prospective Diabetes Study (UKPDS) Outcomes. *Diabetologia*, 47(10):1747–1759, 2004.

[78] Cleve Moler and Charles van Loan. Nineteen Dubious Ways to Compute the Exponential of a Matrix. *SIAM Review*, 20:801–836, 1978.

[79] S. Gallivan, M. Utley, M. Jit, and C. Pagel. A Computational Algorithm Associated with Patient Progressing Modelling. *Computational Management Science*, 4(3):283–299, 2007.

[80] Hendriek C Boshuizen, Stefan K Lhachimi, Pieter Baal, R. T. Hoogenveen, Henriette A. Smit, J. P. Mackenbach, and Wilma J Nusselder. The DYNAMO-HIA model: An efficient implementation of a risk factor / chronic disease Markov model for use in health impact assessment: (accepted for publication). *Demography*, nd.

[81] Evert van Imhoff and Wendy Post. Microsimulation Methods for Population Projection. *Population: An English Selection*, 10(1):97–138, 1998.

[82] Harold Holder, Sven Andreasson, Thor Norström, Esa Österberg, and Ingeborg Rossow. *Estimates of Harm Associated with Changes in Swedish Alcohol Policy*. National Institute of Public Health, Stockholm, March 2005.

[83] H. Bronnum-Hansen. Quantitative health impact assessment modelling. *Scand J Public Health*, 37(5):447–449, 2009.

[84] Linda J. Cobiac, Theo Vos, and Jan J. Barendregt. Cost-Effectiveness of Interventions to Promote Physical Activity: A Modelling Study. *PLoS Med*, 6(7):e1000110, 2009.

[85] C. D. Mathers, Christopher J. L. Murray, and Alan Lopez. Epidemiological evidence: improving validity throught consistency analysis. *Bullentin of World Health Organization*, 80(8):611, 2002.

[86] J. Lennert Veerman, Johan P. Mackenbach, and Jan J. Barendregt. Validity of predictions in health impact assessment. *J Epidemiol Community Health*, 61(4):362–366, 2007.

[87] Jan J. Barendregt, Luc Bonneux, and Paul J. van der Maas. The health care cost of smoking. *The New England Journal of Medicine*, 337(15):1052–1057, 1997.

[88] Eduardo Massad, Marcelo N. Burattini, Luis F. Lopez, and Francisco A. B. Coutinho. Forecasting versus projection models in epidemiology: The case of the SARS epidemics. *Medical Hypotheses*, 65(1):17–22, 2005.

[89] Dennis Bray and Hans von Storch. Prediction or Projection? *Science Communication*, 30(4):534–543, 2009.

[90] Peter Anderson and Bern Baumber. Alcohol in Europe: A public health perspective, 2006.

[91] T. Kelly, W. Yang, C. S. Chen, K. Reynolds, and J. He. Global burden of obesity in 2005 and projections to 2030. *INTERNATIONAL JOURNAL OF OBESITY*, 32(9):1431–1437, 2008.

[92] M. Ezzati and Alan D. Lopez. Smoking and oral tobacco use. In Majid Ezzati, Alan D. Lopez, Anthony Rodgers, and Christopher

J. L. Murray, editors, *Comparative quantification of health risks*, volume 1, pages 883–957. WHO, Geneva, 2004.

[93] Michelle Kruijshaar, Jan Barendregt, and Lonneke de van Poll-Franse. Estimating the prevalence of breast cancer using a disease model: data problems and trends. *Population Health Metrics*, 1(5):1–8, 2003.

[94] C. D. Mathers and D. Loncar. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Medicine*, 3, 2006.

[95] J. I. Levy, J. J. Buonocore, and K. von Stackelberg. Evaluation of the public health impacts of traffic congestion: a health risk assessment. *Environmental health : a global access science source*, 9:65, 2010.

[96] D.F Sullivan. A single index of mortality and morbidity. *HSMHA Health Reports*, 86:347–354, 1971.

[97] Isabelle Soerjomataram, Esther de Vries, Eero Pukkala, and Jan Willem Coebergh. Excess of cancers in Europe: A study of eleven major cancers amenable to lifestyle change. *International Journal of Cancer*, 120(6):1336–1343, 2007.

[98] J. Pomerleau, M. McKee, T. Lobstein, and C. Knai. The burden of disease attributable to nutrition in Europe. *PUBLIC HEALTH NUTRITION*, 6(5):453–461, 2003.

[99] Susan T. Stewart, David M. Cutler, and Allison B. Rosen. Forecasting the Effects of Obesity and Smoking on U.S. Life Expectancy. *N Engl J Med*, 361(23):2252–2260, 2009.

[100] J. Rehm, R. Room, M. Monteiro, G. Gmel, K. Graham, N. Rehn, C. T. Sempos, and D. Jernigan. Alcohol as a risk factor for global burden of disease. *European Addiction Research*, 9(4):157–164, 2003.

[101] David A. Leon and Jim McCambridge. Liver cirrhosis mortality rates in Britain from 1950 to 2002: an analysis of routine data. *Lancet*, 367(9504):52–56, 2006.

[102] Alexander C. Wagenaar, Matthew J. Salois, and Kelli A. Komro. Effects of beverage alcohol price and tax levels on drinking: a meta-analysis of 1003 estimates from 112 studies. *Addiction (Abingdon, England)*, 104(2):179–190, 2009.

[103] C Tigerstedt, T Karlsson, Pia Mäkelä, E Österberg, and I Tuominen. Health in alcohol policies: the European Union and its Nordic Member States. In M. McKee, L. MacLehose, and E. Nolte, editors, *Health policy and European Union enlargement*, pages 111–127. Open University Press, Berkshire, 2004.

[104] Robin C. Purshouse, Petra S. Meier, Alan Brennan, Karl B. Taylor, and Rachid Rafia. Estimated effect of alcohol pricing policies on health and health economic outcomes in England: an epidemiological model. *The Lancet*, 375(9723):1355–1364, 2010.

[105] H. D. Holder and J. O. Blose. Impact of changes in distilled spirits availability on apparent consumption: a time series analysis of liquor-by-the-drink. *British journal of addiction*, 82(6):623–631, 1987.

[106] W. Hollingworth, B. E. Ebel, C. A. McCarty, M. M. Garrison, D. A. Christakis, and F. P. Rivara. Prevention of deaths from harmful drinking in the United States: the potential effects of tax increases and advertising bans on young drinkers. *Journal of studies on alcohol*, 67(2):300–308, 2006.

[107] SHAAP. Alcohol: Price, Policy and Public Health: Report on the Finding of the Expert Workshop on Price convened by Scottish Health Action on Alcohol Problem, 2007.

[108] J. Rehm, J. Klotsche, and J. Patra. Comparative quantification of alcohol exposure as risk factor for global burden of disease. *International journal of methods in psychiatric research*, 16(2):66–76, 2007.

[109] DYNAMO-HIA Consortium. Work Package 8 - Report on data collection for alcohol consumption and related relative risks: available on www.dynamo-hia.eu., 2010.

[110] Alexander C. Wagenaar, Amy L. Tobler, and Kelli A. Komro. Effects of alcohol tax and price policies on morbidity and mortality: a systematic review. *American journal of public health*, 100(11):2270–2278, 2010.

[111] Eurostat. *Eurostat - OECD methodological manual on purchasing power parities*. Office for Official Publications of the European Communities, Luxembourg, 2005.

[112] Philip J. Cook. *Paying the tab: The economics of alcohol policy*. Princeton Univ. Press, Princeton, NJ, 2007.

[113] Samuel H. Preston, Patrick Heuveline, and Michel Guillot. *Demography: Measuring and modeling population processes*. Blackwell, Oxford, 2005.

[114] P. S. Meier, R. Purshouse, and A. Brennan. Policy options for alcohol price regulation: the importance of modelling population heterogeneity. *Addiction (Abingdon, England)*, 105(3):383–393, 2010.

[115] T. Lobstein. Denying obesity? *The European Journal of Public Health*, 19(6):570, 2009.

[116] P. Kopelman, S. A. Jebb, and B. Butland. Executive summary: Foresight 'Tackling Obesities: Future Choices' project. *Obesity reviews : an official journal of the International Association for the Study of Obesity*, 8 Suppl 1:vi–ix, 2007.

[117] J. Lennert Veerman, Eduard F. van Beeck, Jan J. Barendregt, and Johan P. Mackenbach. By how much would limiting TV food advertising reduce childhood obesity? *Eur J Public Health*, 19(4):365–369, 2009.

[118] C. L. Chiang. Competing Risks in Mortality Analysis. *Annu. Rev. Public. Health.*, 12(1):281, 1991.

[119] Susan Griffin, Karl Claxton, Neil Hawkins, and Mark Sculpher. Decision Models Need to be Fit for Purpose for Decision-Making: Response to Caro et al, 2007.

[120] Susan Griffin, Karl Claxton, Neil Hawkins, and Mark Sculpher. Probabilistic Analysis and Computationally Expensive Models: Necessary and Required? *Value in Health*, 9(4):244–252, 2006.

[121] Kenneth J. Rothman and Sander Greenland. *Modern epidemiology*. Lippincott-Raven, Philadelphia, Pa., 1998.

[122] Karl R. Popper. *Unended quest: An intellectual autobiography*. Open Court, La Salle, Ill, 1976.

[123] Rainer Fehr and Odile Mekel, editors. *Quantifying the health impacts of policies – Principles, methods, and models: Scientific Expert Workshop, Düsseldorf, Germany, 16 - 17 March 2010*, volume 11 of *LIGA.Fokus series*. LIGA.NRW, Düsseldorf, 2010.

[124] David Briggs. A framework for integrated environmental health impact assessment of systemic risks. *Environmental Health*, 7(1):61, 2008.

[125] Marko Tainio, Jouni Tuomisto, Otto Hanninen, Juhani Ruuskanen, Matti Jantunen, and Juha Pekkanen. Parameter and model uncertainty in a life-table model for fine particles (PM2.5): a statistical modeling study. *Environmental Health*, 6(1):24, 2007.

[126] Isabelle Soerjomataram, Esther de Vries, Gerda Engholm, Georg Paludan-Müller, Henrik Brønnum-Hansen, Hans H. Storm, and Jan J. Barendregt. Impact of a smoking and alcohol intervention programme on lung and breast cancer incidence in Denmark: An example of dynamic modelling with Prevent. *European Journal of Cancer*, 46(14):2617–2624, 2010.

[127] David L. MacIntosh, Taeko Minegishi, Matthew Kaufman, Brian J. Baker, Joseph G. Allen, Jonathan I. Levy, and Theodore A. Myatt. The benefits of whole-house in-duct air cleaning in reducing exposures to fine particulate matter of outdoor origin: A modeling analysis. *J Expos Sci Environ Epidemiol*, 2009.

[128] Jeremy A. Lauter. *The time dimension in measurments of population health*. PhD thesis, Erasmus University Rotterdam, Rotterdam, Netherlands, 2009.

[129] Johan P. Mackenbach, Irina Stirbu, Albert-Jan R. Roskam, Maartje M. Schaap, Gwenn Menvielle, Mall Leinsalu, and Anton E. Kunst. Socioeconomic inequalities in health in 22 European countries. *The New England journal of medicine*, 358(23):2468–2481, 2008.

[130] DYNAMO-HIA Consortium. Work Package 10 - Report on data collection for cancers: available on www.dynamo-hia.eu., 2010.

[131] DYNAMO-HIA Consortium. Work Package - Report on data collection for COPD: available on www.dynamo-hia.eu., 2010.

[132] DYNAMO-HIA Consortium. Work Package 9 - Report on data collection for CVD, diabetes, stroke and related relative risks: available on www.dynamo-hia.eu., 2010.

[133] DYNAMO-HIA Consortium. Work Package 7 - Report on data collection for obesity and related relative risks: available on www.dynamo-hia.eu., 2010.

[134] DYNAMO-HIA Consortium. Work Package 8 - Report on data collection for alcohol consumption and related relative risks: available on www.dynamo-hia.eu., 2010.

[135] K. Gu, C. C. Cowie, and M. I. Harris. Mortality in adults with and without diabetes in a national cohort of the u.s. population, 1971-1993. *Diabetes care*, 21(7):1138–1145, 1998.

[136] E. Barrett-Connor and K. T. Khaw. Diabetes mellitus: an independent risk factor for stroke? *American journal of epidemiology*, 128(1):116–123, 1988.

[137] A. B. Wirehn, H. M. Karlsson, and J. M. Carstensen. Estimating disease prevalence using a population-based administrative healthcare database. *Scandinavian journal of public health*, 35(4):424–431, 2007.

[138] L. F. Ellison, H. I. Morrison, M. J. de Groh, and P. J. Villeneuve. Health consequences of smoking among Canadian smokers: An update. *Chronic diseases in Canada*, 20(1):36–39, 1999.

[139] Donald R. Shopland and David M. Burns, editors. *Changes in cigarette-related disease risks and their implication for prevention and control*, volume no. 97-4213 of *NIH publication*. National Institutes of Health National Cancer Institute, Bethesda Md, 1997.

[140] MJ Thun, C Day-Lally, DG Myers, and 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 Donald R. Shopland and David M. Burns, editors, *Changes in cigarette-related disease risks and their implication for prevention and control*, volume no. 97-4213 of *NIH publication*, pages 305–382. National Institutes of Health National Cancer Institute, Bethesda Md, 1997.

[141] Peter Tanuseputro, Douglas G. Manuel, Susan E. Schultz, Helen Johansen, and Cameron A. Mustard. Improving population attributable fraction methods: examining smoking-attributable mortality for 87 geographic regions in Canada. *American journal of epidemiology*, 161(8):787–798, 2005.

[142] Kenneth F. Adams, Arthur Schatzkin, Tamara B. Harris, Victor Kipnis, Traci Mouw, Rachel Ballard-Barbash, Albert Hollenbeck, and Michael F. Leitzmann. Overweight, obesity, and mortality in a large prospective cohort of persons 50 to 71 years old. *The New England journal of medicine*, 355(8):763–778, 2006.

[143] Anja Schienkiewitz, Matthias B. Schulze, Kurt Hoffmann, Anja Kroke, and Heiner Boeing. Body mass index history and risk of type 2 diabetes: results from the European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam Study. *The American journal of clinical nutrition*, 84(2):427–433, 2006.

[144] Andrew G. Renehan, Margaret Tyson, Matthias Egger, Richard F. Heller, and Marcel Zwahlen. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. *Lancet*, 371(9612):569–578, 2008.

[145] Christa Meisinger, Angela Döring, Barbara Thorand, Margit Heier, and Hannelore Löwel. Body fat distribution and risk of type 2 diabetes in the general population: are there differences between men and women? The MONICA/KORA Augsburg cohort study. *The American journal of clinical nutrition*, 84(3):483–489, 2006.

[146] James B. Meigs, Peter W. F. Wilson, Caroline S. Fox, Ramachandran S. Vasan, David M. Nathan, Lisa M. Sullivan, and Ralph B. D'Agostino. Body mass index, metabolic syndrome, and risk of type 2 diabetes or cardiovascular disease. *The Journal of clinical endocrinology and metabolism*, 91(8):2906–2912, 2006.

[147] Gary Whitlock, Sarah Lewington, Paul Sherliker, Robert Clarke, Jonathan Emberson, Jim Halsey, Nawab Qizilbash, Rory Collins, and Richard Peto. Body-mass index and cause-specific mortality in 900 000 adults: collaborative analyses of 57 prospective studies. *Lancet*, 373(9669):1083–1096, 2009.

[148] S. C. Moore, S. T. Mayne, B. I. Graubard, A. Schatzkin, D. Albanes, C. Schairer, R. N. Hoover, and M. F. Leitzmann. Past body mass index and risk of mortality among women. *International journal of obesity (2005)*, 32(5):730–739, 2008.

[149] Eugenia E. Calle, Michael J. Thun, Jennifer M. Petrelli, Carmen Rodriguez, and Clark W. Heath. Body-Mass Index and Mortality in a Prospective Cohort of U.S. Adults. *New England Journal of Medicine*, 341(15):1097–1105, 1999.

[150] J. Stevens, J. Cai, E. R. Pamuk, D. F. Williamson, M. J. Thun, and J. L. Wood.

The effect of age on the association between body-mass index and mortality. *The New England journal of medicine*, 338(1):1–7, 1998.

[151] Gabriela Vazquez, Sue Duval, David R. Jacobs, and Karri Silventoinen. Comparison of body mass index, waist circumference, and waist/hip ratio in predicting incident diabetes: a meta-analysis. *Epidemiologic reviews*, 29:115–128, 2007.

[152] UK National Audit Office. Tackling obesity in England: London. London.

[153] Daniel L. McGee. Body mass index and mortality: a meta-analysis based on person-level data from twenty-six observational studies. *Annals of epidemiology*, 15(2):87–97, 2005.

[154] A. Bergström, P. Pisani, V. Tenet, A. Wolk, and H. O. Adami. Overweight as an avoidable cause of cancer in Europe. *International journal of cancer. Journal international du cancer*, 91(3):421–430, 2001.

[155] J. R. Banegas, F. Rodríguez-Artalejo, A. Graciani, F. Villar, and R. Herruzo. Mortality attributable to cardiovascular risk factors in Spain. *European journal of clinical nutrition*, 57 Suppl 1:S18–21, 2003.

[156] Rik P. Bogers, Wanda J. E. Bemelmans, Rudolf T Hoogenveen, Hendriek C Boshuizen, Mark Woodward, Paul Knekt, Rob M. van Dam, Frank B. Hu, Tommy L. S. Visscher, Alessandro Menotti, Roland J. Thorpe, Konrad Jamrozik, Susanna Calling, Bjørn Heine Strand, and Martin J. Shipley. Association of overweight with increased risk of coronary heart disease partly independent of blood pressure and cholesterol levels: a meta-analysis of 21 cohort studies including more than 300 000 persons. *Archives of internal medicine*, 167(16):1720–1728, 2007.

[157] Tone Bjørge, Steinar Tretli, and Anders Engeland. Relation of height and body mass index to renal cell carcinoma in two million Norwegian men and women. *American journal of epidemiology*, 160(12):1168–1176, 2004.

[158] Daphne P. Guh, Wei Zhang, Nick Bansback, Zubin Amarsi, C. Laird Birmingham, and Aslam H. Anis. The incidence of co-morbidities related to obesity and overweight: a systematic review and meta-analysis. *BMC public health*, 9:88, 2009.

[159] D. Michal Freedman, Alice J. Sigurdson, Preetha Rajaraman, Michele M. Doody, Martha S. Linet, and Elaine Ron. The mortality risk of smoking and obesity combined. *American journal of preventive medicine*, 31(5):355–362, 2006.

[160] Debbie A. Lawlor, Carole L. Hart, David J. Hole, and George Davey Smith. Reverse causality and confounding and the associations of overweight and obesity with mortality. *Obesity (Silver Spring, Md.)*, 14(12):2294–2304, 2006.

[161] Julia Hippisley-Cox, Carol Coupland, John Robson, Aziz Sheikh, and Peter Brindle. Predicting risk of type 2 diabetes in England and Wales: prospective derivation and validation of QDScore. *BMJ (Clinical research ed.)*, 338:b880, 2009.

[162] Zhengming Chen, Gonghuan Yang, Maigeng Zhou, Margaret Smith, Alison Offer, Jieming Ma, Lijun Wang, Hongchao Pan, Gary Whitlock, Rory Collins, Shiru Niu, and Richard Peto. Body mass index and mortality from ischaemic heart disease in a lean population: 10 year prospective study of 220,000 adult men. *International journal of epidemiology*, 35(1):141–150, 2006.

[163] V. J. Carey, E. E. Walters, G. A. Colditz, C. G. Solomon, W. C. Willett, B. A. Rosner, F. E. Speizer, and J. E. Manson. Body fat distribution and risk of non-insulin-dependent diabetes mellitus in women. The Nurses' Health Study. *American journal of epidemiology*, 145(7):614–619, 1997.

[164] Anders Engeland, Steinar Tretli, and Tone Bjørge. Height and body mass index in relation to esophageal cancer; 23-year follow-up of two million Norwegian men and women. *Cancer causes & control : CCC*, 15(8):837–843, 2004.

[165] R. Nyamdorj, Q. Qiao, T. H. Lam, J. Tuomilehto, S. Y. Ho, J. Pitkäniemi, T. Nakagami, V. Mohan, E. D. Janus, and S. R. G. Ferreira. BMI compared with central obesity indicators in relation to diabetes and hypertension in Asians. *Obesity (Silver Spring, Md.)*, 16(7):1622–1635, 2008.

[166] World Cancer Research Fund / American Institute for Cancer Research. *Food, nutrition, physical activity and the prevention of cancer: A global perspective : a project of World Cancer Research Fund International*. AIC, Washington, D.C, 2007.

[167] Gillian K. Reeves, Kirstin Pirie, Valerie Beral, Jane Green, Elizabeth Spencer, and Diana Bull. Cancer incidence and mortality in relation to body mass index in the Million Women Study: cohort study. *BMJ (Clinical research ed.)*, 335(7630):1134, 2007.

[168] A. Rosengren, H. Wedel, and L. Wilhelmsen. Body weight and weight gain during adult life in men in relation to coronary heart disease and mortality. A prospective population study. *European heart journal*, 20(4):269–277, 1999.

[169] J. W. Yarnell, C. C. Patterson, H. F. Thomas, and P. M. Sweetnam. Comparison of weight in middle age, weight at 18 years, and weight change between, in predicting subsequent 14 year mortality and coronary events: Caerphilly Prospective Study. *Journal of epidemiology and community health*, 54(5):344–348, 2000.

[170] Youfa Wang, Eric B. Rimm, Meir J. Stampfer, Walter C. Willett, and Frank B. Hu. Comparison of abdominal adiposity and overall obesity in predicting risk of type 2 diabetes among men. *The American journal of clinical nutrition*, 81(3):555–563, 2005.

[171] International Agency for Research on Cancer (IARC). *Alcohol Consumption and Ethyl Carbamate: Meeting summary*, volume 96 of *IARC Monographs on the Evaluation of Carcinogenic Risks to Human*. IARC, Lyon, 2007.

[172] C. Tabak, H. A. Smit, L. Räsänen, F. Fidanza, A. Menotti, A. Nissinen, E. J. Feskens, D. Heederik, and D. Kromhout. Alcohol consumption in relation to 20-year copd mortality and pulmonary function in middle-aged men from three european countries. *Epidemiology (Cambridge, Mass.)*, 12(2):239–245, 2001.

[173] I R White, D R Altmann, and K Nanchahal. 'optimal' levels of alcohol consumption for men and women at different ages, and the all-cause mortality attributable to drinking, 2000.

[174] Jürgen Rehm, Urszula Sulkowska, Marta Mańczuk, Paolo Boffetta, John Powles, Svetlana Popova, and Witold Zatoński. Alcohol accounts for a high proportion of premature mortality in central and eastern europe. *International journal of epidemiology*, 36(2):458–467, 2007.

[175] Courtney C. Greene, Katharine A. Bradley, Chris L. Bryson, David K. Blough, Laura E. Evans, Edmonds M. Udris, and David H. Au. The association between alcohol consumption and risk of copd exacerbation in a veteran population. *Chest*, 134(4):761–767, 2008.

[176] Ian R. White, Dan R. Altmann, and Kiran Nanchahal. Alcohol consumption and mortality: modelling risks for men and women at different ages. *Bmj*, 325(7357):191, 2002.

Thanx!

Umwege erhöhen die Ortskenntnis.

Already after my first degree, in political science, I realized I wanted to try myself at a thesis. Little I knew and much less I understood that becoming a scientist is much more a process of socialization than of conscious learning. It is developing the – more often than not – intuitive understanding of when and to whom a given argument is considered acceptable, innovative, or just odd. I was fortunate to go through this process at two distinctive places, the *Erasmus MC* and the *National Institute for Public Health*. Both are places of excellence in their own right and being exposed to diverse opinions only fostered my development. I am deeply grateful to both institutions and the individuals therein that helped and supported me in this – not always carefree – process.

As the path to this thesis was certainly not a linear one, in a way starting in Berlin and leading via sojourn in the US and the Baltic Sea to an EU-research project in Rotterdam, there are so many to thank, so many to mention. I am not even trying to be as exhaustive as it would be necessary. Those who feel left out wrongly, should make their grievances known and can hope for a drink in return; in many ways a more tangible expression of gratitude than the mentioning in a publication of arguably limited entertainment value. (Nobody working at NS needs to apply.)

At the *Max-Planck-Institute for Demographic Research*, James W. Vaupel who so generously sponsored my stay at Duke. Gerda Neyer, as head of the policy lab, for being a honest and helpful senior researcher. Alyson van Raalte, not only trusted proofreader and friend, but also later colleague at MGZ. Francesco Lagona for introducing me to spatial statistic. Sandra Krapf and Tobias Kluesner, former future colleagues at the policy lab, for many pleasant lunch breaks. At the *Sanford School for Public Policy, Duke University*, I enjoyed a two-year stay that certainly changed me in many ways. In particular the classes by Arthur "Art" Spengler turned me into a more effective writer and taught me to work diligently yet cheerful on any given problem. Phil Cook enabled my participation in real research project and was unconditional supportive when supervising my master thesis. Helen 'Sunny' Ladd for being a generous host on many occasions. Kurt, Kelvin, Jesse, Ervin, Elke, and more who became friends while solving the world's problems (one memo at a time). At the *Institute for Statistics and Econometrics* Wolfgang Härdle, Marlen Müller, and Peter Wilrich for being inspiring statisticians and instructors.

At *EMI (Expertise Centre for Methodology)*, my department at the *RIVM*, I always felt like a full member of this fine team. Fondly I remember the sport en speldag and many coffee breaks with the statistics boys Albert Wong, Jan de Chateau, Jose Ferreira, and Maarten Schippers (whose desk and chair I sequestered while being there), Pieter the Baal, Rudolf Hoogenveen, Arnold, and Rene. Thanks for the advice, cookies, and fun chats.

Hendrik Boshuizen started as a co-promoter but ended up being my full-promoter by becoming a professor sometime after I became her PhD student.[1] I am deeply indebted to her as I certainly wouldn't have finished this thesis without her gentle and constant support. One of her former PhD students described her once to me, rightfully, as a very noble person. I can only second that and might add that she is also impressively good at what she is doing. Furthermore, I can assure future PhD students of hers that she suffers fools gladly and has a great sense of humor.

At *MGZ*, Lennert Veerman to whom a special thank must go for being one of the driving forces behind the DYNAMO-HIA grant application. Frank van Lenthe who was

[1] I clearly see no causation in this correlation, but then who can truly tell...

a mentor to me on several occasions. My colleagues at MGZ, in particular the international kids club: Lifang Liu (the good soul of our part of the hallway) , James O'Mahoney (beer buddy and sparing partner), Margarete Kulik, Elizabeth Wever (landlord of last resort), Mauricio Avendano and Isabelle Soerjomataram (actual landlords, office-mates, and friends), Matejka Rebolik, Rasmus Hoffmann, Alyson van Raalte (again), Pieter Baal (for the second time), Istvan Majer, and the other German in zee department: Stefan.

Wilma Nusselder, my co-promoter, who gave me such a warm welcome when I started at MGZ, hard working, and always a true colleague. This thesis would look much different without her, especially her attention to detail was always helpful.

Johan Mackenbach, my promoter, head of the department, and outstanding researcher. Always prepared, always to the point, and, unfortunately showing it far too seldom, a slightly mischievous sense of humor. Always generous in supporting the research of his PhD students and improving their research findings.

At *iMTA*, Hans Severens, Maiwenn Al, Ken Redkop, and Laura Burger for being a great team and being very understanding that I still had to finish parts of my thesis in the first weeks of my new job. Hedwig, Parida, Maria, and Pieter (for the third time) for many chatty lunch breaks.

At *various* or – God forbid – *without any Institutional Affiliations*, Daniela for providing a port of refuge in nearby Brussels whenever I longed for classicistic architecture, French food in German quantities, and Belgian beer. Thomas, Birte, Kjell, and Jona for being cute-to-look-at at the breakfast table. Dr. Dominik Nagl who through many long hours and short-communications was omnipresent via Skype and, luckily, had an almost counter-cyclical mood and motivation-level while working on his thesis, allowing us to cheer each other up most of the times. Sabine, also a virtual companion during these testing times, was always diligent proof reader of first, last, and at times only resort, who just recently couldn't be helped and started her own "process of socialization". Last but certainly not least, I have to thank my mother who always let me wander my own path and, equally, my grandmother who always gave me the feeling to be special.

And everybody who gave me a smile every now and then...