RIVM report 260706001/2005

A conceptual framework for budget allocation in the RIVM Chronic Disease Model. A case study of Diabetes mellitus.

RT Hoogenveen, TL Feenstra, PHM van Baal,
CA Baan

Contact: TL Feenstra
Department for Prevention and
Health Services Research (PZO)
Talitha.Feenstra@rivm.nl

This investigation has been performed by order and for the account of the RIVM, within the framework of project S/260706, Priority setting in chronic diseases: methodology for budget allocation.

RIVM, P.O. Box 1, 3720 BA Bilthoven, telephone: 31 - 30 - 274 91 11; telefax: 31 - 30 - 274 29 71

Rapport in het kort

Conceptueel model voor budgetallocatie met het RIVM Chronische Ziekten Model. Toepassing bij Diabetes mellitus.

Dit rapport beschrijft de elementen van een zogeheten ‘budget allocatie model’. Dit model is bedoeld ter ondersteuning van beleidsmakers bij keuzes over de inzet van budget voor primaire preventie en/of preventie in de zorg bij chronische aandoeningen. Als concrete toepassing is gekozen voor Diabetes mellitus.

Een uitbreiding van het RIVM Chronische Ziekten Model beschrijft het verband tussen diabetes, risicofactoren en hart- en vaatziektecomplicaties. Een gezondheidseconomische module berekent vervolgens gezondheidseffecten in termen van gewonnen levensjaren en voor kwaliteit van leven gecorrigeerde gewonnen levensjaren (QALYs), interventiekosten, en kosten van zorg. Ten slotte bespreken we hoe de voorkeuren van beleidsmakers kunnen worden geformaliseerd in doelstellingsfuncties en (budget-)beperkingen.

Deze drie elementen zijn de basis voor een toepassing van budgetallocatie bij diabetes. De ontwikkelde methode is ook toepasbaar bij andere chronische ziekten, omdat we het bredere RIVM Chronische Ziekten model als uitgangspunt hebben gebruikt. Het nieuwe model voor diabetes is niet alleen een basis voor budgetallocatie, maar ook op zichzelf al bruikbaar om primaire preventie en verschillende vormen van preventie van complicaties bij diabetes te evalueren. Het model kan voor deze interventies de consequenties voor Nederland berekenen, zowel voor de kosten van zorg als voor de gezondheid.

Abstract

A conceptual framework for budget allocation in the RIVM Chronic Disease Model. A case study of Diabetes mellitus.

The research project 'Priority setting in chronic diseases: methodology for budget allocation' aims to develop a methodology to support optimal allocation of the health care budget with respect to chronic diseases.

The current report describes the modelling steps required to address budget allocation questions regarding the prevention of chronic diseases and their complications with the RIVM Chronic Disease Model, with specific attention to diabetes mellitus.

An extension of the RIVM Chronic Disease Model deals with the links between diabetes, its risk factors and its macrovascular complications. A health economics module computes outcomes in terms of intervention costs, costs of care and composite health effects. Finally, it is discussed how to formalize different preferences of policy makers in various objective functions and constraints.

These three elements form the basis for the analysis of budget allocation questions in diabetes care. The model allows for the comparison of primary prevention with the prevention of complications in diagnosed patients as to costs of care and health effects. Furthermore, as it stands, the model with the health economics module per se is a useful tool for policy analysis, for instance, to compare the costs and effects of different interventions.

Voorwoord

Het MAP SOR-onderzoeksprogramma ‘Methodologie optimale gezondheidswinst en kwaliteit van zorg’ wijst op het strategisch belang van methodeontwikkeling ter ondersteuning van optimale inzet van het gezondheidszorgbudget. Het project ‘Budgetallocatie: methode-ontwikkeling voor prioritering van interventies bij chronische ziekten (Priority setting in chronic diseases: methodology for budget allocation)’ sluit daarbij aan, en beoogt bij te dragen aan de methodes voor het vergelijken van de kosten en effecten voor preventieve screenings en zorginterventies. Daartoe zal een zogeheten budgetallocatiemodel worden ontwikkeld. Het onderzoeksproject richt zich daarbij vooral op die budgetallocatieproblemen die net een stap verder gaan dan traditionele kosteneffectiviteitsanalyses en bijvoorbeeld afwegingen maken tussen verschillende types interventies voor een ziekte. Er wordt gestreefd naar methodologie die voor verschillende (chronische) aandoeningen toepasbaar is. Daarbij is voor diabetes gekozen als eerste voorbeeldstudie.
Het voorliggende rapport beschrijft de conceptuele en formele opzet van een diabetes model wat geschikt is als basis voor budgetallocatie (hoofstukken 2 en 3). Daarnaast wordt de benadering voor budgetallocatie conceptueel uitgewerkt (hoofdstuk 4).
Het onderzoek voor de hoofdstukken over het diabetesmodel in dit rapport is uitgevoerd in nauwe samenwerking met de onderzoekers die waren betrokken bij de beantwoording van de kennisvraag ‘Preventie van diabetes’. Het werk aan en de rapportage voor beide projecten is op elkaar afgestemd, om dubbelingen te voorkomen. Daarbij zullen de inhoudelijke onderdelen worden gerapporteerd in het rapport ‘Modelling Chronic Diseases: the diabetes module. Justification of (new) input data in the mathematical RIVM Model’, terwijl in dit rapport de conceptuele en formele opzet van het model is beschreven. Dit onderzoeksproject bouwt gedeeltelijk voort op eerder werk over het onderwerp budgetallocatie in het kader van het ZONMW programma doelmatigheid. De gezondheidseconomische module die een van de bouwstenen vormt van het budgetallocatiemodel is beschreven in het rapport ‘Cost Effectiveness Analysis with the RIVM Chronic Disease Model’. Tenslotte willen we Hendriek Boshuizen en Guus Den Hollander bedanken voor nuttig commentaar bij conceptversies van het rapport.

Contents

Summary

Introduction
The research project 'Priority setting in chronic diseases: methodology for budget allocation' is aimed at developing a methodology to support optimal allocation of the health care budget with respect to chronic diseases. Diabetes is an interesting example, because many different options for primary prevention and the prevention of complications exist.

Objective
The aim of the current report was to describe the modelling steps required to address budget allocation questions for the prevention of chronic diseases, with specific attention for diabetes.

Methods
The scope of the budget allocation problem was limited to either the case of a single disease, or the case of a single type of interventions. Based on the RIVM chronic disease model, a multistate transition model was developed with joint states representing individuals' risk factor and disease status. Specific attention was paid to the modelling of diabetes, our example for the single disease case. A health economics module enabled the computation of total costs, intervention costs, total effects and cost-effectiveness ratios. Finally, the conceptual approach to budget allocation was developed, using mathematical programming to formalize the problem, and paying attention to objective functions and budget constraints.

Results
The result of the modelling efforts is a set of formal equations defining the elements relevant for diabetes in the RIVM Chronic Disease Model 2005 joint version and a health economics module. The implementation of these in mathematica has to be combined with estimates of the input data. Then, the model is ready for the evaluation of different prevention interventions for diabetes and its macrovascular complications. Budget allocation problems can then be addressed by adding an objective function and formulating the relevant constraints. The resulting optimisation problem may be solved either in a single step or in a two step procedure, first generating results from scenario analysis of the model and then optimizing over these outcomes.

Conclusion
The implemented model will enable us to evaluate both interventions for the primary prevention of diabetes and interventions in diabetes patients to prevent macrovascular complications. This forms the basis for the analysis of budget allocation questions in diabetes care. The general structure of our budget allocation model is not limited to diabetes and is intended to be applicable for any chronic disease. The disease model per se is a useful tool to give policy relevant information in combination with the health economics module, for instance, the costs and effects of different interventions can be compared.

1. Introduction

1.1 Background

The research project 'Priority setting in chronic diseases: methodology for budget allocation' is aimed at developing a methodology to support optimal allocation of the health care budget with respect to chronic diseases. A so-called budget allocation model will be set up, this is an explicit model of the objectives and constraints faced by decision makers in their choices between different health care interventions. The explicit formulation of the model helps to analyse several issues that complicate matters for the decision maker. Thus, we hope to make the issue of budget allocation more transparent and insightful for decision makers. In this introductory chapter the economic and epidemiological background for this project are described. First, the problem of budget allocation and its relation with cost effectiveness analysis is explained. Second, the scope of the budget allocation model for this project is described. The next section introduces diabetes mellitus, which was the case study for this project. The chapter ends with a description of the report's aim and contents.

1.2 Cost effectiveness analysis and budget allocation

Economic evaluation has been developed as a tool to inform policy makers about the costs and effects of medical interventions to support their decisions on the optimal allocation of health care resources. Usually cost effectiveness analyses compare the outcomes of two or more alternatives and result in a cost-effectiveness ratio, which expresses how much money has to be paid per additional unit of health gained (for instance, life years (LY's) or quality adjusted life years (QALY's) gained). The lower this ratio, the more cost effective it is to implement the investigated intervention. That is, the more health effects are obtained for given expenditures. Some interventions turn out to be dominant, because they are less costly and, at the same time, generate more health effects than their comparator. Other interventions result in better health but at additional costs.

To support decision makers in allocating money to different interventions in health care, a cost effectiveness ratio alone, though useful, may not be sufficient. In addition, decision makers may also need information that more explicitly addresses the issue of budget allocation and for instance compares the total costs and effects of interventions (so called budget impact analyses), or even explicitly formulates the objectives and constraints in a budget allocation model. To estimate the total effects and costs of interventions over time often more epidemiological data and demographic data is needed than in ordinary cost effectiveness analyses. For instance, the number of patients requiring certain treatments is needed to compute the total costs of the intervention, while a cost effectiveness analysis can be based on the costs per patient. Consider a decision maker that wants to allocate his budget over different treatments for the same chronic disease. Assume that he has the choice between supporting a (tertiary) prevention program targeted at a large group of patients in mild stages of the disease, with a cost effectiveness ratio of 31000 per LY gained or supporting a new

surgical procedure with exactly the same cost effectiveness ratio, but targeted at patients in the more advanced stages of the disease. That is, the net present values of the incremental costs and effects of both interventions are equal. Of course, they differ widely in the distribution of these costs and effects over time. They may also differ in their budgetary consequences. Assuming that the total health effects and costs of different interventions are available, the next question is how the decision maker should go on to compare the efficiency of these interventions? If the prevention program effectively limits the number of patients in need of surgical procedure, then it is clear that the total costs and effects of the interventions considered are interdependent. Such interdependencies require a model where time enters explicitly.

Budget allocation models combine the results of cost effectiveness analysis with epidemiological and demographic data, an optimality criterion, and budget constraints to find the optimal allocation of resources over programs. The best-known budget allocation model has maximization of the sum of health effects as its optimality criterion, under the constraint that the sum of program costs remains within a given total budget.[1] We refer to this model as the standard model. The current report describes the modelling steps required to address budget allocation questions for the prevention of chronic diseases and their complications, with specific attention for diabetes mellitus.

1.3 Scope of the budget allocation model

It is our aim to develop a methodology to support optimal allocation of the health care budget with respect to chronic diseases that enables one to compare the costs and health effects of primary prevention with secondary prevention (screening) and tertiary prevention (prevention of complications in diagnosed patients). To enable this, the first step in the development of the budget allocation model is to ensure that different interventions can be compared in terms of costs and effects. The comparison of health costs and effects and cost effectiveness ratio's from different studies is surrounded by difficulties because of differences in adopted methodologies, perspective and differences in data sources.

To develop a method to consistently deal with health effects and costs on the population level we used the RIVM Chronic Disease Model (CDM) [2]. This is a multistate transition model that links prevalence of risk factors to the incidence of 28 chronic diseases. The model allows to compute the effects of a reduction in risk factors through prevention on life years gained and quality of life, taking account of comorbidity. The CDM models the entire Dutch population, following the life course of birth cohorts over time and thus allows to estimate the total costs and health effects of interventions for the entire Dutch population. By using the CDM to compute health effects and costs it can be ensured that the same methodology and the same type of costs and effects are taken into account for all different interventions that are compared [3]. The RIVM Chronic Disease Model was intended primarily as tool to model the health effects of primary prevention. Since our aim is to also compare primary with secondary and tertiary prevention and to use economic as well as health outcomes, the CDM must be adjusted in several ways. As a guide for these

adjustments in the CDM and in the development of the budget allocation model diabetes was chosen as a case study.

1.4 A case study of diabetes mellitus

Diabetes mellitus is a major source of morbidity and mortality, associated with serious complications, loss in quality of life and high use of health care.[4 5] In the recent decades, the incidence and prevalence of diabetes have increased and it was estimated that the number of diabetic patients in the Netherlands will increase with 36% in 2020 due to demographic trends only. [6 7] It is to be expected that this increase will be even larger, considering the observed trends in obesity and physical inactivity.[8] The present health care system is not optimally organized for an adequate treatment and control of chronic diseases like type 2 diabetes, the health care budget for prevention is limited, and the burden for health care providers is high. Therefore, it is of great importance to know which type of prevention strategy for diabetes provides the largest gain in terms of health (quality of life, life years gained) and in terms of savings in health care use relative to its input requirements.

The severity and prevalence of the disease make diabetes a candidate for various prevention strategies. However, to be able to quantify more accurately the choice of strategy, requires knowledge on the (side-)effects of the possible strategies. Three types of prevention strategies of diabetes exist: primary, secondary and tertiary prevention. The aim of primary prevention is to prevent the development of diabetes in high risk individuals in the general population. Several risk factors of diabetes have been identified, with overweight and lack of physical activity being the major ones. The aim of secondary prevention is the early detection and subsequently treatment of patients with yet undiagnosed diabetes. About half of patients with diabetes is yet undiagnosed.[6] Tertiary prevention aims at obtaining health gains by the delay or even prevention of complications as a result of intensive follow-up and treatment of diagnosed diabetes patients. Several clinical trials demonstrated that a good glycemic control as well as adequate control of lower leg morphology, and treatment of risk factors for cardiovascular diseases (hypertension, dyslipidemia, overweight, smoking) can considerably limit the incidence of diabetes complications.[9 10]

To obtain reliable estimates for the Dutch situation, it is necessary to analyze and compare different interventions in the same setting, using one model and comparable epidemiological outcome and cost data for all interventions. For primary prevention of diabetes this has been done by the iMTA in collaboration with the WHO.[11 12] Many economic evaluations of single or combined strategies for the prevention of macrovascular complications exist. [13 14] Earnshaw[15] evaluated different combinations of four interventions, all aiming at the prevention of complications in diabetes patients and used budget allocation methods to find optimal allocations for different objectives and constraints. He used the CDC diabetes model, representing the USA situation as a basis. In our case study, we want to compare both interventions for primary prevention and for the prevention of macrovascular complications.

1.5 Overall aim of the project and content of the current report

In short, this research project should result in budget allocation models that inform decision makers on the optimal allocation for different objectives or constraints. The strict assumptions of the standard model involve a fixed patient group, a single budget constraint, and a static model without attention for the distribution of costs and effects over time. This standard model underlies the decision rules of cost effectiveness analyis. By choosing diabetes as a case study, we hope to tackle a lot of the methodological issues that can be encountered if one wants to compare primary prevention with secondary and tertiary prevention. The general structure of our budget allocation model is not limited to diabetes and is intended to be applicable for any chronic disease. In a second case, different interventions in primary prevention (smoking cessation and interventions to reduce overweight) will be compared.

To set up the budget allocation model for the diabetes case study, the following research questions have to be answered :

- What are the characteristics of a diabetes model that can be applied for budget allocation?
- What data are available to populate a model presenting the Dutch situation?
- What additional modeling is needed to actually perform a budget allocation analysis?
- What methodological issues arise and which are relevant to approach?

To answer these questions, the following topics will be addressed in the current report. To characterize a diabetes model suitable for budget allocation, first, a review is given of existing diabetes models (section 2). Second, the structure of a diabetes model that is suitable for budget allocation is described (section 2), starting with a description of the current diabetes model in the RIVM Chronic Disease Model. Third, the formal approach to the diabetes model is outlined (section 3) and the data requirements to estimate model parameters are described (section 3). The estimates of these parameters will be reported in this report's twin report for the diabetes model.[16]. Fourth, section 4 shortly describes the health economics module,[16] contains a description of the aims of budget allocation, and introduces the additional requirements to perform a budget allocation analysis based on this model together with the methodological issues to be addressed. Finally section 5 concludes with a summary and discussion of the results.

2. A model for evaluation of interventions for the case of diabetes

The model was based on the existing diabetes module in the RIVM Chronic Disease Model.[2] However, this model had to be updated and extended to allow for explicit modelling of diabetes complications in relation to risk factors. This section starts with a description of diabetes models in the literature and what can be learned from then, followed by an explanation of the changes made on the Chronic Disease Model (CDM).

2.1 Review of existing models in the literature

The modelling work in the project was started with a scan of the literature on existing diabetes models, to put our work into perspective and to see whether any models suitable for budget allocation already existed. A summary of the review is given in appendix A.

From an examination of the characteristics of the models identified, three observations may be made. First, the number of different models is quite limited, because many models were adjustments of others. For microvascular complications, one standard structure exists, for instance described in publications on the UKPDS model. [17 18] Mortality was only increased for the most severe microvascular complication stages (blindness, ESRD, amputation). For macrovascular complications, most models were based on the Framingham risk functions,[19] sometimes with added modelling to enable for recurrent events. Second, for microvasculair complications, the most important explanatory variables used in the models were diabetes duration and HbA1c, while for macrovascular complications these were age, sex, SBP, cholesterol ratio, and smoking status. Third, most models described stochastic individual life courses using discrete time steps of 1 year. Few models were time continuous and deterministic, but these did not include duration as an explanatory variable.

Important sources for parameter estimates used in many models were the Diabetes Control and Complications Trial (DCCT) [20], the United Kingdom Prsopective Diabetes Study (UKPDS) [21], the Wisconsin Epidemiologic Study of Diabetic Retinopathy[22], and the Framingham Heart Study.[19] The risk functions of the latter were often used for the cardiovascular complications. Simulation models have been developed based on each large follow-up study, such as UKPDS and DCCT.

To conclude, most models focussed on microvascular complications and modelled macrovascular complications in less detail. From a budget allocation point of view, the macrovascular complications deserve more attentions, since they account for about 40% of total costs of care related to diabetes, while microvascular complications account for about 10%. [23] Two models with extensive modeling of macrovascular complications were the UKPDS-model [17] and the Swiss model by Palmer et al..[24-26]

2.2 Model structure

This subsection describes the general set up of a chronic disease model with diabetes suitable for budget allocation over all types of prevention.

2.2.1 The RIVM chronic disease model

The current model (CDM2003) has been described previously.[2 27]

In short, the RIVM Chronic Disease Model (CDM) has been developed as a tool to describe the morbidity and mortality effects of autonomous changes of and interventions on chronic disease risk factors taking into account integrative aspects. The model contains the following risk factors: cholesterol, systolic blood pressure, smoking, physical activity level, and Body Mass Index. It models 28 chronic diseases: cardiovascular diseases, distinguishing acute myocardial infarction, other coronary heart disease, stroke, and chronic heart failure, COPD, asthma, diabetes mellitus, dementia, several musculoskeletal disorders, and 15 different forms of cancer.
The mathematical model structure is called a multi-state transition model and is based on the life table method. The model states defined are the risk factor classes and disease states. State transitions are possible due to changes between classes for any risk factor, incidence, remission and progress for any disease, and mortality. The model describes the life course of cohorts in terms of changes between risk factor classes and changes between disease states over the simulation time period. Risk factors and diseases are linked through relative risks on disease incidence. That is, incidence rates for each risk factor class are found as relative risks times baseline incidence.
The main model parameters are:

- the initial population numbers
- initial class prevalence rates and transition rates for all risk factors
- initial prevalence, incidence, remission and mortality rates for all diseases and
- relative risk values specified by risk factor and chronic disease

All model parameters and variables are specified by gender and age. The time step used for modeling is 1 year.
The main model outcome variables are incidence, prevalence and mortality numbers specified by disease, and integrative measures such as total and quality-adjusted life years.

The CDM2003 model describes risk factor class prevalence and disease prevalence numbers separately. For example, the model keeps track of the number of smokers and non-smokers, and the number of persons without and with diabetes, but not of the number of persons with diabetes who smoke. It takes account of the dependency relations between risk factors and diseases through a time-dependent covariance matrix. The CDM2005 joint version which will be presented below describes the joint prevalence numbers explicitly. For example, the number of non-smokers without diabetes, non-smokers with diabetes etcetera.

2.2.2 A diabetes model suitable for budget allocation

Diabetes is one of the 28 chronic diseases which is explicitly modeled in CDM.
To allow for analysis of the complications of diabetes, diabetes is modeled both as a disease and a risk factor for some other diseases, that is, an intermediate disease.
Some cardiovascular diseases are also intermediate diseases in CDM. Figure 1 shows the causal dependency structure between diabetes and cardiovascular diseases.

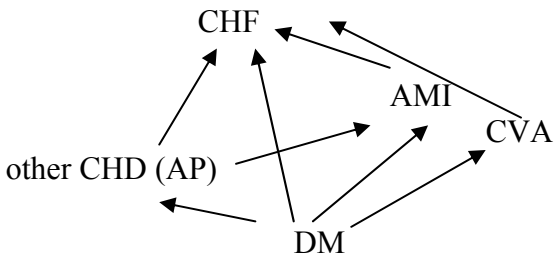


Figure 1: Causal dependency relations between diabetes mellitus and several cardiovascular diseases[i]

[i] CHF=Chronic heart failure, AMI= Acute Myocardial infarction, other CHD=other Coronary diseases (AP=Angina pectoris), CVA=Stroke (cardiovascular accident), DM=Diabetes Mellitus

For each pair of diseases the incidence rate ratios of the 'end' disease are adjusted for incidence through the 'intermediate' disease (see section 3.3.1.2).

The following risk factors included in the current model are important for the modelling of diabetes and its complications:

- Body Mass Index (BMI)
- physical activity
- smoking
- total cholesterol
- Systolic Bloodpressure (SBP)

For all risk factors, the model distinguishes several classes, for instance normal weight (BMI<25), overweight (BMI 25-30) and obese BMI (BMI >30).
In CDM2003, BMI and physical activity are risk factors for diabetes incidence as well as for some other cardiovascular diseases. Smoking, cholesterol and blood pressure are modeled as risk factors for cardiovascular diseases.
Adding these risk factors for diabetes (BMI and activity) and its complications to the figure above, a rather complex structure results.

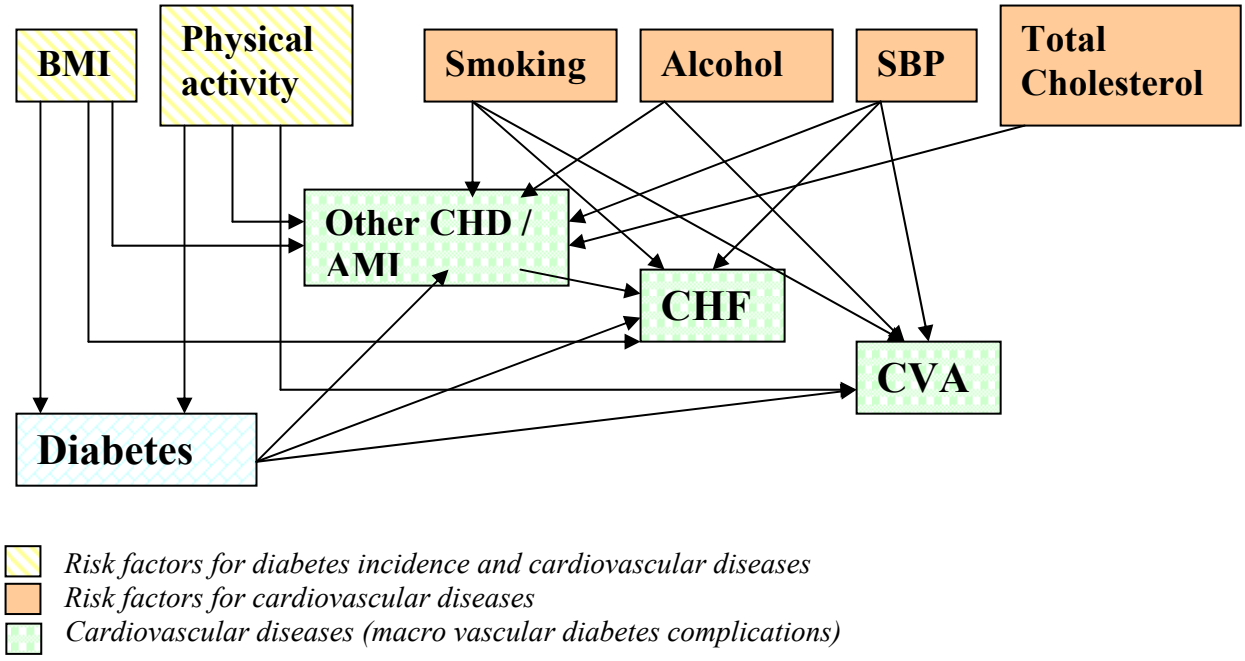


Figure 2: Structure of dependency relations between risk factors, diabetes mellitus and several cardiovascular diseases.

Simplifying this structure (see Figure 3), the CDM2003 marginal model catches the link between risk factors, diabetes and its complications, but does not keep track of the risk factors for people with diabetes.

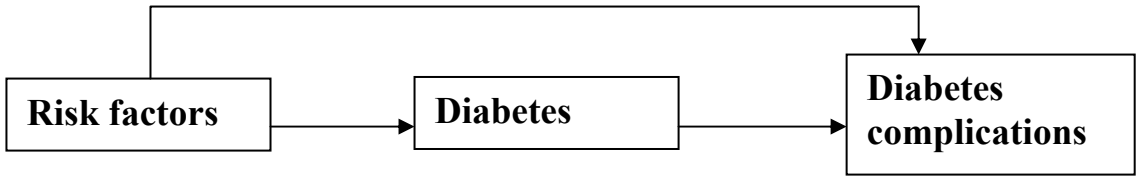


Figure 3: Simplified structure of dependency relations between risk factors, diabetes mellitus and complications in CDM-2003

Therefore, the model is fit to evaluate the effects of primary prevention, since the effect of changes in risk factor prevalences on diabetes and on complications can be analysed. However diabetes is modeled as a single stage disease and in the model the risk of complications is not different for a person with diabetes with or without e.g. high blood pressure. Therefore, the model is not fit to evaluate the effect of tertiary prevention, that is the prevention of diabetes complications resulting from improved care. Hence, an extension of the model is needed to allow for the evaluation of tertiary prevention.
To be able to evaluate prevention of complications, the model was to be extended to include the prevalence of risk factors for macrovascular complications in diabetes, as follows (see Figure 4):

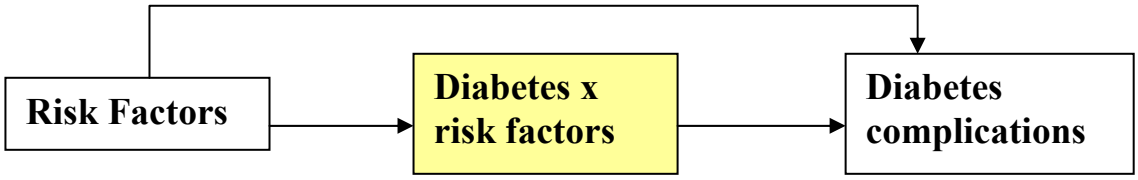


Figure 4: Simplified structure of dependency relations between risk factors, diabetes mellitus and complications in CDM2005

That is, in the new model (CDM2005 joint version), the diabetes population was divided into risk factor classes. This enables us to evaluate the effect of treatment aiming at risk factor levels in patients with diabetes to reduce the incidence of macrovascular complications. For the formal model, this new structure implied that the model had to be reformulated, keeping track of risk factor prevalences, once people get a disease. The evaluation of the effects of treatment on the level of glycemic control, asks for a further extension. People with diabetes then have to be classified according to their level of HbA1c.

3. A model of diabetes and macrovasculair complications in relation to risk factors.

This section describes the formal structure and the input data needed for the diabetes module and related parts in the recent version of the RVM Chronic Disease model named CDM2005 joint version. This version was developed to enable budget allocation over all types of prevention for diabetes and its macrovascular complications. The general structure of the model was described in section 2 above. The current section starts in 3.1. with a list of the methodological issues to be addressed in realizing this structure followed by an explanation of the elements in the RIVM Chronic Disease Model 2005. Then, 3.2 discusses the input parameters needed and refers to the documentation of their estimates from empirical data. Finally, section 3.3 addresses each of the methodological issues, in a description of the joint model set up.

3.1 The RIVM Chronic Disease model, CDM2005 joint version

The RIVM Chronic Disease Model (CDM2005 joint version) refers to the combination of a conceptual model, mathematical formulas that specify the approaches to the methodological issues involved in realizing the conceptual model and the implementation of the mathematical formulas. Implementation involves transcription of the mathematical formulas in a Mathematica code and the estimation of model parameters based on empirical data.
Many methodological issues that arise are not specific to diabetes and were solved for the chronic disease model in general. For some issues, reference is given to the relevant background reports to keep this report as specific as possible.

To realize a formal model according to the structure set out in section 2 above, the following issues were addressed:
- the modelling of socalled intermediate diseases, diseases that are a risk factor for other diseases in the model
- attribution of mortality to diabetes and its complications, that is, adjustment of disease-related excess mortality rates for competing mortality risks.
- joint modelling of risk factor prevalence and disease prevalence
- modelling of the effects of disease duration on transition rates
- stages of diabetes and modelling of progression over these stages

The RIVM Chronic Disease Model has been developed as a mathematical tool to describe the relation between any selected set of risk factors and set of chronic diseases over time. Therefore, methodological requirements for any model version are flexibility and internal consistency. That is, the model must work for any selection of risk factors and diseases, while results for any disease or risk factor should not change in the first order with the selection of the remaining risk factors and diseases. Moreover, since the number of different model states grows exponentially with the number of risk factors and diseases included, the model implementation must be efficient.[16]

The CDM is a Markov-type multistate transition model. This model type provides a mathematically and statistically consistent way of dealing with disease risks that are dependent through joint risk factors and that depend on both time and age. The term multistate means that persons belong to one of a set of disjoint states that are characterized by the values of the state variables. Since our model applies to chronic diseases the states are defined in terms of risk factor classes and disease conditions. The term transition means that any change of state over time was modeled through transitions between these states. The term Markov-type means that the future states are independent on the past states conditional on the current state. In other words, all relevant information to describe the future life course of the cohort is stored in the current values of the state variables selected.
The model consists of an initialisation part and a simulation part. The initial input data have to be corrected and combined to calculate the input variables for the model simulations. This happens in the model initialization part. In the model simulation part the time-continuous changes of the model variables are calculated, i.e. in case of the joint model the changes of all state prevalence numbers. This part consists of differential equations that describe the change of the model state prevalence rates over time.

In Figure 5 and Figure 6 we present the general state-transition structure for any risk factor and disease respectively.

Figure 5 presents the state-transitions for any risk factor. The horizontal arrows describe the transitions between the risk factor classes for those that survive, the vertical arrows describe the transition to the state ‘deceased’, i.e. mortality. For instance, for the risk factor smoking, the classes are never smokers, smokers and former smokers and n equals 3. Arrows from class 1 to class 2 and class 2 to class 2 represent smoking initialization and smoking cessation, an arrow from class 3 to class 2 represents relapse, while from class 2 back to class 1 no arrow exists, since returning to the class of never smokers is not possible. In the figure only transitions between neighboring classes were shown for graphical simplicity, but this is no restriction to the model. In case of treatment, additional classes and transitions may be introduced.

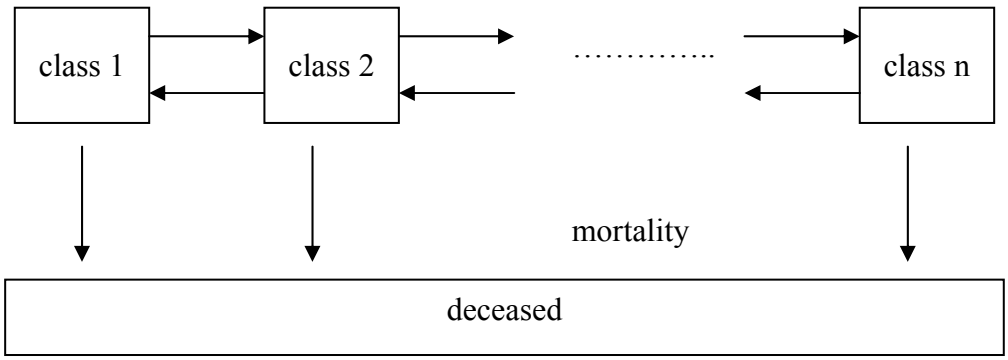


Figure 5: Transitions between risk factor classes and mortality

Figure 6 presents the state-transition structure for any disease. We distinguished only one disease condition here, i.e. with the disease. In principle, the mathematical model

can deal with any number of disease states. The horizontal arrows describe the transitions between the states, i.e. disease incidence from the state ‘without the disease’ to ‘with the disease’. The vertical arrows describe the transition to the state ‘deceased’, i.e. mortality. Remission is the possibility to return from a disease state to a state without the disease. Disease incidence and as a result also disease mortality rates depend on the risk factor prevalences.

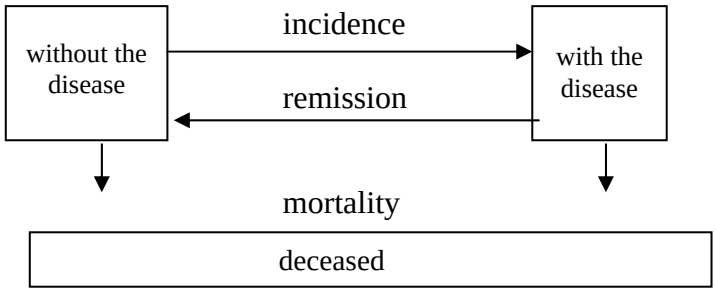


Figure 6: Transitions between disease states and mortality

Summarizing, Figure 7 shows all possible transitions in a time step Δt with ncr(r) denoting the number of risk factor classes for risk factor r.

time t	small change in time Δt	time t+Δt
gender g	no changes	≡ g
age a	aging= Δt	= a + Δt
risicofactor r classes 1 to ncr(r)	transitions between classes	 classes 1 to ncr(r)
disease d not present present	transitions between states ↓ ↑	 not present present
alive yes no	mortality ↓	 yes no

Figure 7: Summary of possible transitions

The main assumptions that were used to formalize the joint CDM model versions are:

(1) We assumed the individual life courses to be statistically independent conditional on all covariates, i.e. the risk factor classes and disease conditions. In other words, all individuals behave independently given their epidemiological characteristics.
(2) We assumed that all risk factors and chronic diseases have a discrete distribution. For continuous risk factors cut-off points were introduced to define a set of disjoint classes.
(3) We assumed the initial risk factor class prevalence rates independent. The initial disease prevalence rates were assumed independent conditional on these risk factors and on causally related intermediate diseases ('local independence assumption'). In the CDM2003 marginal model version the class transition rates for each risk factor were also assumed to be independent from the other risk factors and the disease states. The CDM2005 joint version allowed us to condition the class transition rates on the disease state. For example, new cases of other CHD among current smokers may have higher cessation rates than CHD-free current smokers.
(4) The incidence rates for each disease were assumed to be independent from the other diseases, conditional on the risk factor levels and the causally related intermediate diseases ('local independence assumption'). For example, the incidence rates for CVA and for other CHD were assumed independent conditional on the epidemiologic risk factors and the diabetes state. We included the most important joint risk factors in the model to cover the dependency relations between these diseases.
(5) We assumed that for any disease the excess mortality is independent from the risk factor levels. This means that the risk factors affect the disease prognosis only through increased risks for other diseases and mortality from other causes of death.
(6) We assumed that the disease-specific excess mortality rates are additive. The latter assumption of additive excess mortality rates was based on the additive cause-specific mortality hazard model.[28]
(7) We assumed multiplicative incidence and mortality risks, i.e. with no interaction on the log-linear scale. For example, the risk for incidence of Diabetes for an individual with both a high BMI and a low activity level is the multiplication of the risk rates for BMI and activity.
(8) The other causes mortality rates depend on the risk factors, but conditional on the risk factors do not depend on the disease states. That is, the risk to die of an accident is the same for a person with or without diabetes, conditional on risk factors.
(9) The risk rate for the incidence of diseases is an approximation for the risk rate for the prevalence of diseases. This assumption is used to derive some of the formulas below.

3.2 The new model of diabetes and macrovasculair complications in relation to risk factors: input data

Here, the type of input data used is shortly described. The following inputdata were estimated, with inputdata new for CDM2005 joint version in bold:

- For each disease, prevalences in the model startyear, incidence rates and remission (for example for asthma) and mortality rates.
- For each risk factor, its division into classes, prevalences in the model startyear for each risk factor class, and transition rates between these classes (e.g. start and stop smoking).
- For each combination of a risk factor with a disease, per risk factor class the relative risks for incidence of the disease.
- For each combination of a causal disease with an 'end' disease, relative risks for incidence of the 'end' disease.
- **Start prevalences in a diabetes population, for each risk factor distinguished**
- **Transition rates between classes in a diabetes population for each risk factor distinguished,**
- **Relative risks for risk factors in a diabetes population for incidence of macrovascular diseases.**

The report 'Modelling Chronic Diseases: the diabetes module. Justification of (new) input data in the mathematical RIVM Model' [29] documents updates of the data that were already present in CDM2003, as well as the estimates used for the new model parameters listed above. All data must be age- and sex specific. Transition, prevalence, incidence and mortality data need to apply to the Dutch population, and were therefore based on the most appropriate, recent Dutch registry data. Relations between risk factors and diseases (relative risks) were assumed to be less country specific and estimated from the international literature.

3.2.1 Population numbers

Demographic input data in CDM are from Statistics Netherlands. [30] The most important input data are the Dutch population numbers in the start year, divided into age and gender, prognosed numbers of births in each year, prognosed migration numbers, and all cause mortality specific to age and gender.

3.2.2 Prevalence of DM and riskfactors in startyear

In the joint model, prevalence has to be specified according to risk factors for complications. These joint prevalences can be computed based on relative risks and total prevalences (see 3.3.1.6). However, for diabetes, specific empirical data were also gathered for the most important risk factors, to validate the model computations. These empirical data were estimated for each risk factor separately, because the estimation of joint classess of all risk factors would result in unreliable estimates.

The importance of the inclusion of a riskfactors follows from:
- its impact on complications
- available evidence on interventions for this risk factor
- prevalence of the risk factor in the Dutch DM population

For the following risk factors additional empirical data were gathered :
- total cholesterol
- systolic Bloodpressure
- body Mass Index
- physical activity
- smoking
- level of HbA1c

For all these risk factors except HbA1c, both the prevalence in the total population as well the prevalence in the DM population was estimated. For HbA1c, only the prevalence in the DM population was estimated.

3.2.3 Transitions in a year

DM incidence

Total diabetes incidence was modelled as the weighted average of the different risk factor and disease classes using relative risks (see section 3.3.1.7). After incidence in DM, the new cases also have to be distributed over the DM risk factor classes.

For the risk factors that affect the incidence of diabetes, the relative risks for incidence of diabetes used in the model will influence the distribution of diabetes incidence over the risk factor classes. This is the case for:
- body mass index
- physical activity
- smoking

For the risk factors that do not affect DM incidence, incidence in DM will be according to the prevalence in the total population. An adjustmentstep (see 3.3.1.7) is needed to adjust this incidence to the prevalence distribution of the risk factor in DM. This is the case for:
- total cholesterol
- systolic Bloodpressure

Finally, the level of HbA1c is not modelled in the total population, mainly due to a lack of data, and hence the level of glycemic control in the incidence of diabetes will be an input parameter and was estimated from empirical data.

Transitions between risk factor classes.

This refers to transitions from e.g. low bloodpressure to higher bloodpressure, for the following risk factors:
- total cholesterol
- systolic Bloodpressure
- body Mass Index
- physical activity
- smoking

In the CDM2005 joint version, transition rates between risk factor classes for a certain risk factor were assumed equal for all combinations of joint states that reflect a transition between these classes. That is, the smoking cessation rate is the same in people with and without diabetes and with a high or a low Body Mass Index. From empirical data, transition rates for specific groups, for example, for diabetes patients could be estimated. Furthermore, if interventions affect these transition rates, they will differ between groups that receive an intervention and those without. Therefore, the model was structured such that transition rates can be adjusted for each specific joint state. For HbA1c, the transition rates were estimated for the diabetes population.

Mortality
Unadjusted excess mortality rates were estimated for DM, and its complications: AMI, other CHD, CVA and CHF. The excess mortality rates were then adjusted as described in section 3.3.1.4, based on the comorbidity rates as computed in section 3.3.1.3.

3.2.4 Relative risks
The incidence of the DM complications (AMI, other CHD, CVA and CHF) in the DM population as well as the incidence of these diseases in the general population without DM is linked to risk factor classes trough relative risks. These relative risks were adjusted for other, confounding factors included in the model to prevent double counting see sections 3.3.1.21 and 2.
Total incidence of complications was estimated from empirical data. Relative risks of the risk factors BMI and activity on DM incidence were used to distribute total incidence over different risk factor classes in the general population.
Furthermore, relative risks for incidence of AMI, other CHD, CHF and CVA in the general population were estimated, as well as the relative risks of intermediate diseases on ‘end’ diseases, for instance, the risk of diabetes on AMI.

3.2.5 Costs of care for each disease stage
An estimate of total costs of diabetes care has to include the costs related to complications. We estimated Dutch costs for diabetes, based on data from the 1999 Cost of Illness in the Netherlands.

However, to analyze the effect of interventions on costs of care, it is relevant to distinguish between cost of care for diabetes alone and costs related to complications which may be prevented by the interventions. To obtain this, we excluded the costs of modeled complications from the total cost estimate and used the model instead, including costs of care for cardiovascular diseases to estimate these costs directly. Because the same relative risk estimates were used in the cost of diabetes estimate and in the model, for a current practice scenario, the model computes the same total diabetes costs.
If hypothetical scenarios would be computed with free interventions which reduce rates of cardiovascular complications, then total diabetes costs will decrease. Of course, interventions are not free, so that this decrease must then be corrected for the costs of the intervention.
Ideally we also want to divide the costs of diabetes without costs of cardiovascular complications over the different classes of HbA1c level. This will then allow

analyzing interventions which result in better control of the level of HbA1c for their consequences on costs of care for diabetes.

3.2.6 Quality of life in each disease stage

The Chronic Disease Model uses estimates from the Global Burden of Disease Study and its Dutch Counterpart [4 31-34] to attribute loss of quality of life weights to diseases. For Diabetes, these weights will be further specified for the different classes of HbA1c level. The effect of the presence of macro vascular complications is modeled through comorbidity. To compute the quality of life in joint states with different diseases present, an assumption has to be made about how comorbidity affects quality of life. Three possible assumptions are that the quality of life weight for a state with both disease A and B could be 1) the lowest of the weights for A or B, 2) the multiplication of the two weights, or 3) the addition of the two weights. These three assumptions were analyzed and their results were compared. This is reported in detail in the report on the health economics module.[16] In the CDM2005 joint version, it is possible to choose among the three different assumptions.

3.3 Methodological issues in a diabetes model, formal solutions in CDM2005 joint version

To start the formal description of the model structure, the variables available as input data in the model will be listed. Let d= A,B,C,D,… denote the nd diseases in the model. Let r=R,S,… denote the nrd risk factors in the model, each with risk factor classes i=1,..,ncr(r). All variables used were age and gender specific, but for ease of notation, these arguments will be ignored below. Then, input data for the model are:

em_d	disease d excess mortality rate
m_{tot}	all cause mortality rates
p_d	disease d prevalence rates
inc_d	disease d incidence rates
cf_d	disease d case fatality rate, i.e. the 1-month mortality rate after disease onset
rem_d	disease d remission rates
p^r_i	class i prevalence rate for risk factor r
λ^r_{ij}	transition rates between risk factor classes, for each risk factor r and all classes i,j
$RR^r_{i,tot}$	the relative risk for total mortality for risk class i of risk factor r
$RR^r_{d,i}$ (unadj)	the relative risk in risk class i of risk factor r for incidence of disease d, unadjusted for intermediate diseases
$RR_{d,A}$ (unadj)	the disease A relative risk for incidence of disease d unadjusted for intermediate diseases

The values for the transition rates and prevalence rates were estimated from empirical data, while the relative risks were estimated based on analyses of data obtained from literature reviews. For diabetes, more detailed estimates were available on risk factor prevalence and on the relative risks for incidence of complications.[29]

Given the input variables, in the model initialization steps the variables needed at the start of the simulation are calculated.

3.3.1 Model initialization steps

At model initialization, the following steps are taken:

1. adjust incidence risk rates of one disease on another for intermediate diseases
2. adjust incidence risk rates of risk factors on diseases for intermediate diseases
3. calculate co-morbidity prevalence rates
4. adjust excess mortality rates for double-counting mortality numbers
5. calculate mortality for other causes
6. compute prevalence rates in joint risk factor and disease classes
7. calculate incidence rates for all model states

3.3.1.1 Adjustment of co-morbidity disease incidence risks for intermediate diseases

For each pair of successively causally related diseases we adjusted the incidence risk rates of one disease on another for the intermediate disease. For example (see Figure 6), the relative risks of DM on CHF are adjusted for other CHD being an intermediate disease between DM and CHF.

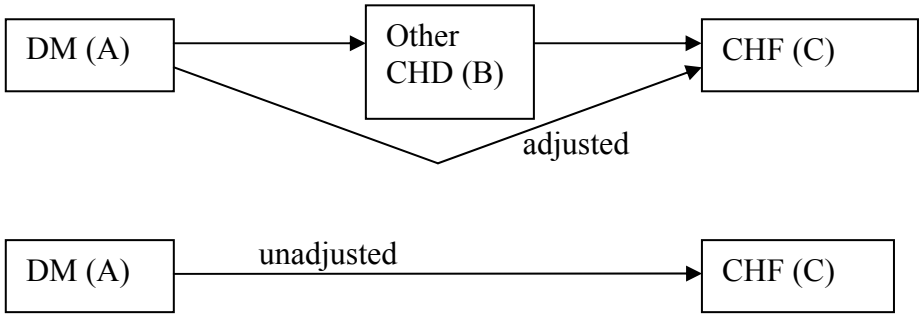


Figure 8: Example of complication risks also working through intermediate disease

We derived the unadjusted relative risk of DM on CHF by rewriting the unadjusted risk rate, conditioning on the intermediate disease other CHD, i.e. as a function of the adjusted relative risk (see appendix B). The resulting formula was then rewritten to express the adjusted relative risk as the unadjusted risk times an adjustment factor:

$$RR_{C,A} = RR_{C,A}(unadj) *$$

$$\{1+(RR_{B,A} - 1)\, p_A+(RR_{C,B} - 1)\, p_B\}/\{1+(RR_{B,A} -1)\, p_A+(RR_{C,B} -1)RR_{B,A}\, p_B\}$$

with:

$RR_{C,A}$	the disease C relative risk for disease A, that is the relative risk for individuals with A to get C, adjusted for intermediate disease B.
$RR_{C,A}(unadj)$	the disease C relative risk for disease A unadjusted for intermediate disease B
$RR_{B,A}$	the disease B relative risk for disease A
$RR_{C,B}$	the disease C relative risk for disease B

p_A, p_B the prevalence rate for disease A and B respectively

The formula shows that the adjustment depends on the prevalence rates of both the causal disease (A) and the intermediate disease (B), and the relative risks of all causal relations involved. The smaller the disease B prevalence rate, the smaller the adjustment.
In the extreme case of $p_B = 0$, the formula reads $RR_{C,A} = RR_{C,A}(unadj)*\{1+(RR_{B,A}–1)\ p_A\}/\{1+(RR_{B,A} –1)\ p_A\}= RR_{C,A}(unadj)$ and the adjustment has no effect. Another example is the case that disease A is no risk factor for B ($RR_{B,A} = 1$). Then, the formula reads: $RR_{C,A} =RR_{C,A}(unadj) *\{1 + (RR_{C,B} –1)\ p_B\}/\{1 +(RR_{C,B} – 1)*1*p_B\}= RR_{C,A}(unadj)$.

For more complex relations between diseases, like those shown in Figure 1, the adjustment has to take place for all relations respectively.[16] This calculation method is possible as long as the causally related disease pairs constitute a so called directed a-cyclic graph. That means, if one disease works as an intermediate disease for another, the latter one is no intermediate disease for the former. Assessing the a-cyclic graph of causally related diseases is a graph-theoretical problem, and has been solved using the so-called adjacency matrix of causally related disease pairs. That is, referring to Figure 1, all dependency relations around diabetes were formalized in a square matrix with a separate row/column for each disease. This matrix contains a ‘1’ if the disease in the column is caused by the disease in the row and a zero otherwise. Then the matrix was reordered so that it formed an upper diagonal matrix. From this matrix, the order of the dependency relations forming an acyclic graph could be obtained. For the structure in Figure 1 the following ordered list of pairs of diseases results:
{{DM, other CHD},{other CHD, AMI},{AMI,CHF},{DM, AMI},{other CHD,CHF} ,{DM,CHF}}.
The adjustment is done going trough this list from the beginning to the end. That is, first the RR of DM on AMI is adjusted for other CHD, then the RR of other CHD on CHF is adjusted for AMI, and finally the RR of DM on CHF is adjusted for AMI.

3.3.1.2 Adjusting disease incidence risks of risk factors for intermediate diseases

For each pair of causally related diseases the epidemiological risk factors affect the dependent ‘end’ disease in two ways, directly as well as indirectly through the independent ‘causal’ disease. For example, (see Figure 9) the dependent ‘end’ disease other CHD incidence risk depends on the prevalence rate of the intermediate disease DM for any risk factor class of the risk factor BMI.

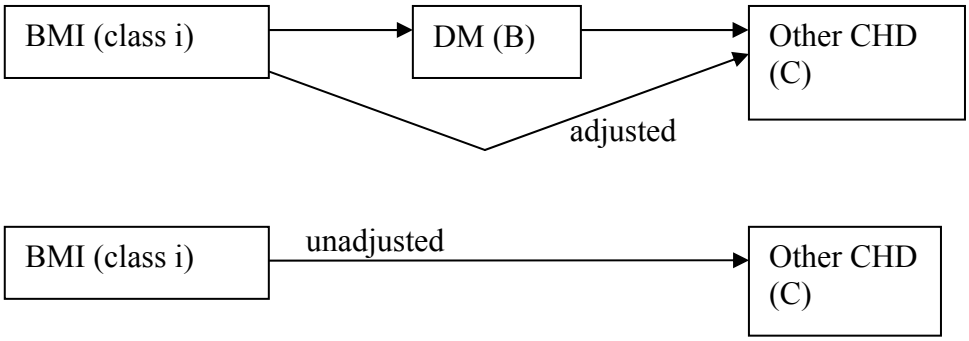


Figure 9: Adjustment of (other CHD) relative risks for risk factors for intermediate disease (DM)

Note the similarity to Figure 6 above. Analogously to the adjustment of relative risks of one disease on another, the unadjusted relative risk can be rewritten as the adjusted relative risk times a factor (see appendix B). Rewriting this formula then expresses the adjusted relative risk as the unadjusted risk times an adjustment factor:

$$RR^{r}_{C,i} = RR^{r}_{C,i}(unadj) *$$

$$\{ E(RR^{r}_{B}) + (RR_{C,B} - 1)\, p_B \} / \{ 1 + (RR_{C,B} - 1)\, (RR^{r}_{B,i} - 1)\, p_B \}$$

with:

$RR^{r}_{C,i}$	disease C incidence relative risk for risk factor class i adjusted for intermediate disease B
$RR^{r}_{C,i}(unadj)$	disease C incidence relative risk for risk factor class i unadjusted for intermediate disease B
$RR^{r}_{B,i}$	disease B incidence relative risk for risk factor class i
$E(RR^{r}_{B})$	mean value of disease B relative risk over all classes of risk factor r
$RR_{C,B}$	disease C incidence relative risk for disease B adjusted for risk factor
p_B	disease B prevalence rate

This formula is analogous to the formula in subsection 3.3.1.1. To see this, note that that for diseases, the $E(RR_B)$, that is, the mean value of $RR_{B,A}$ over all classes (states) of the causal disease A equals $p_A * RR_{B,A} + (1 - p_A) * 1 = (RR_{B,A} - 1) * p_A + 1$.

Again, the calculation method is possible as long as the causally related disease pairs constitute a directed a-cyclic graph.[16] The formula was then applied recursively, that is going from the beginning to the end along the list of all ordered pairs of related diseases. For example, at first the relative risks of for instance the risk factor BMI on other CHD were adjusted for the intermediate disease DM, then those of BMI on AMI for the intermediate diseases DM and other CHD, etcetera.

3.3.1.3 Calculation of initial disease co-morbidity rates

Using the relative risks mentioned in the foregoing paragraph, we calculated the initial prevalence rates for all co-morbidity disease states. The calculation steps applied were the following.[35]

(1) We calculated the co-morbidity prevalence risks resulting from joint risk factors using the relative risk values for these risk factors. For example, smoking as a joint risk factor results in co-morbidity of other CHD and lung cancer. We used the assumptions of independently distributed risk factors and multiplicative risks:

$$p_{A,B} = \Pi_r \Sigma_i RR^r_{A,i} RR^r_{B,i} p^r_i \; p_A p_B / (\Pi_r \Sigma_i RR^r_{A,i} p^r_i * \Pi_r \Sigma_i RR^r_{B,i} p^r_i)$$

with:

p_A , p_B	prevalence rate of disease A and B respectively
$p_{A,B}$	joint prevalence (= co-morbidity) rate of both diseases A and B
$RR^r_{A,i}$, $RR^r_{B,i}$	relative risk of disease A and B respectively for class i of risk factor r
p^r_i	prevalence rate of class i of risk factor r

The formula shows how the disease co-morbidity rate depends on the risk factors through the relative risks. If one disease is not associated with any risk factor (the relative risks have value 1 for all r), the co-morbidity rate is equal to the product of the disease prevalence rates and the diseases are independently distributed.

(2) We adjusted the co-morbidity prevalence risks for all causally related disease pairs using the empirically known disease incidence risks adjusted for intermediate diseases (see 3.3.1.1). For example, DM as an independent risk factor for CHF results in co-morbidity between DM and CHF.

(3) We adjusted the co-morbidity prevalence risks for joint intermediate diseases. For example, co-morbidity between DM and CHF also results from the effect of DM on CHF through the intermediate disease AMI.

(4) We adjusted the co-morbidity prevalence risks for joint ‘causal’ diseases. For example, co-morbidity between AMI and CHF results from other CHD (Angina Pectoris) being an independent risk factor for both AMI and CHF.

(5) Finally, since all co-morbidity risks were defined as prevalence rate ratios we calculated for each causally related pair of diseases the co-morbidity rate ratio for the reversed pair. For example, the prevalence rate ratio for other CHD given disease CHF was calculated from the prevalence rate ratio for CHF given other CHD.

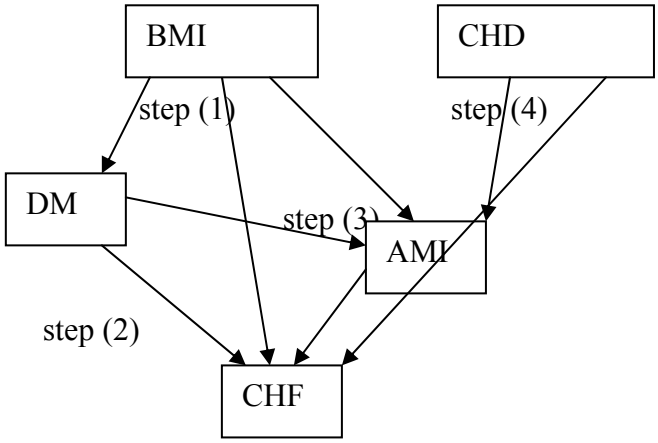


Figure 10: Example of steps to calculate co-morbidity rates

These calculation steps result in co-morbidity prevalence rate ratios for each pair of diseases that are related through joint risk factors and/or directly or indirectly through intermediate diseases along a pathways of causal relations.

If empirical data are available on the comorbidity prevalence rates then an alternative approach is to use these empirical estimates.

3.3.1.4 Adjustment of excess mortality rates

CDM distinguishes two types of disease-related mortality: (1) mortality with the disease, and (2) mortality due to the disease. The model parameter related to (1) is the excess mortality rate, i.e. the excess mortality rate of persons with the disease compared to those without the disease. The excess mortality rates can also be caused by other co-morbid chronic diseases such as disease complications, e.g. other CHD being a complication of diabetes. Summing up excess mortality numbers over all diseases then results in double-counting mortality cases. Therefore we also need (2), the adjusted disease-related excess mortality rates. These excess mortality rates are adjusted for competing death risks, and thus can be interpreted as the mortality uniquely attributable to the disease. Unadjusted disease-related excess mortality rates can be estimated from empirical data.[29 36 37] We calculated the excess mortality rates from empirical disease incidence and prevalence rates available from registries in general practice using the DisMod model equations.[38] The DisMod equations express the relation between disease incidence, prevalence and mortality rates. They allow computing one of these rates from data on the other two.
Figure 11 presents all cause mortality, disease-related mortality, and mortality from other causes of death in relation to co-morbidity.

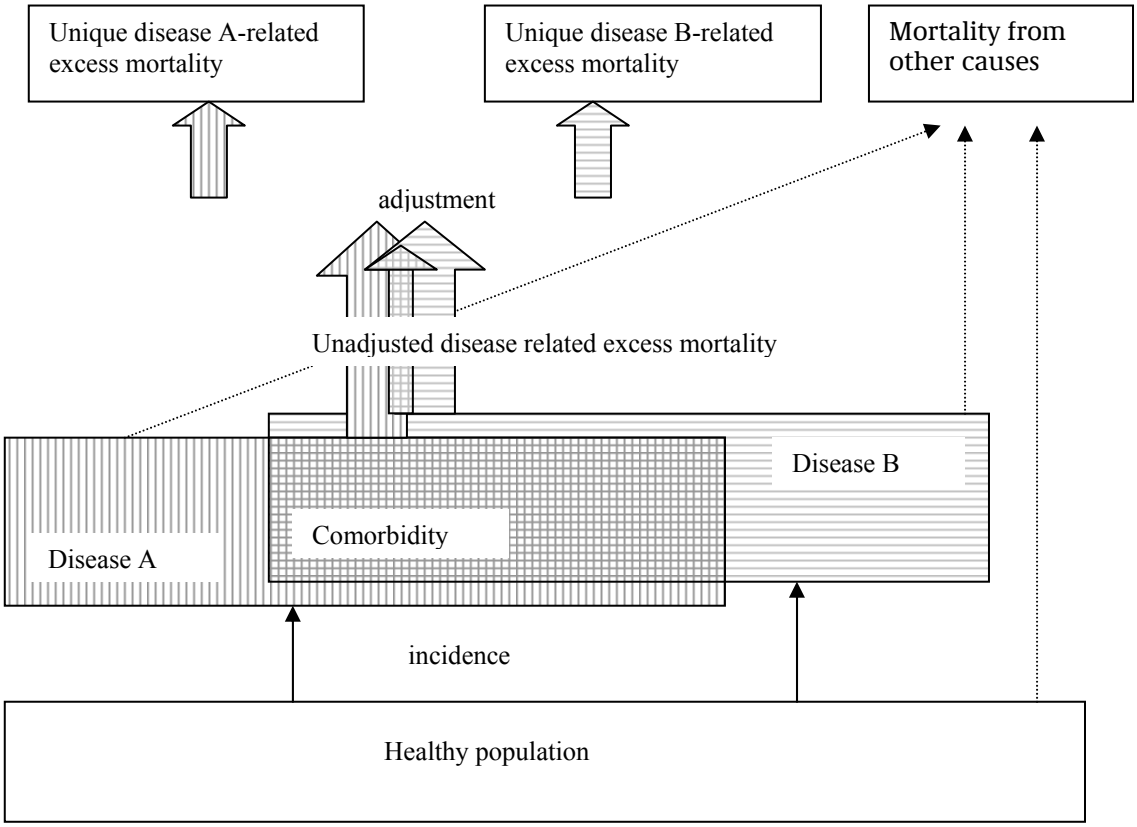


Figure 11: Relation between all cause and disease-related mortality

The co-morbidity prevalence rates were calculated in the foregoing paragraph (3.3.1.3). Using these co-morbidity prevalence rates we related the unadjusted disease-related excess mortality to the adjusted ones (see appendix C for details)[39]:

$$em_A = em_{A0} + \{ \Sigma_{d \neq A} (p_{d,A} - p_A p_d)/p_A(1-p_A) \} em_{A0}$$

with:

em_A, em_{A0}	disease A excess mortality rate unadjusted and adjusted for co-morbidity respectively
p_d, p_A	disease d and disease A prevalence rate
$p_{d,A}$	joint prevalence (= co-morbidity) rate of diseases d and A

The term ($p_{d,A}$ - p_A p_d) is the calculated co-morbidity prevalence rate minus the co-morbidity rate found in case of independent disease rates. The term $p_A(1-p_A)$ is used to scale the expression and equals $p_{A,A}$ - p_A p_A , the 'co-morbidity of a disease with itself' (note that $p_{A,A}= p_A$). Combining these equations for all diseases result in a matrix equation:

$$diag(M) \{em\} = M \{em_0\}$$

with:

M	matrix with elements ($p_{d,A}$ - p_A p_d).
diag(M)	the diagonal-matrix of M , that is, the matrix with elements $p_A(1- p_A)$
	{em}, $\{em_0\}$ vector of excess mortality rates unadjusted and adjusted respectively

The matrix equation was solved:

$$\{em_0\} = M^{-1} \text{ diag}(M) \{em\}$$

In this way the empirically known unadjusted excess mortality rates were adjusted for double-counting mortality numbers through co-morbidity.

3.3.1.5 Mortality rates and rate ratios for other causes

The all cause mortality rates are the sum of the adjusted excess mortality rates and the mortality rates from other causes of death. All cause mortality numbers are available from Statistics Netherlands (CBS). The adjusted disease-related excess mortality rates were calculated above. We combined both to calculate the mortality rates for other causes of death. We took account of the 'acute' mortality of diseases, i.e. the 1-month mortality (case fatality) rate immediately after disease onset. This type of mortality was modeled only for diseases with unstable initial disease periods, i.e. AMI and CVA:

$$m_{oc} = m_{tot} - \Sigma_d \, em_{d0} \, p_d - \Sigma_d \, inc_d \, cf_d$$

with:

m_{oc}	mortality rate for other causes of death
m_{tot}	all cause mortality rate
p_d	disease d prevalence rates
em_{d0}	disease d excess mortality rates adjusted for co-morbidity
inc_d	disease d incidence rate
cf_d	disease d case fatality rate, i.e. the 1-month mortality rate after disease onset

Likewise we calculated the relative risks for other causes of death, i.e. the relative mortality risks through all diseases that were not included in our model. The equations are expressed in so called risk multipliers, these are re-scaled relative risks, obtained by dividing the relative risk with the weighted average of relative risks over all risk classes.[40]

$$RM^r_{i,oc} = \{ RM^r_{i,tot} \, m_{tot} - \Sigma_d \, RM^r_{d,i} \, [em_{d0} \, p_d + inc_d \, cf_d \,] \} \, / \, m_{oc}$$

with the same notation as above and:

$RM^r_{i,oc} = RR^r_{i,oc} / \Sigma_i \, RR^r_{i,oc} \, p^r_i$	risk multiplier for other causes of death for class i of risk factor r
$RM^r_{i,tot} = RR^r_{i,tot} / \Sigma_i \, RR^r_{i,tot} \, p^r_i$	risk multiplier for all cause mortality
$RM^r_{d,i} = RR^r_{d,i} / \Sigma_i \, RR^r_{d,i} \, p^r_i$	risk multiplier for disease d incidence in risk class i of risk factor r.
p^r_i	class i prevalence rate for risk factor r

These equations may be solved for the $RR^r_{i,oc}$.

3.3.1.6 Initial prevalence rates for all model states

Next the initial prevalence rates for all joint states have to be calculated. That is, sthe prevalence rates for each value of the vector (i,j,…nrd,A,B,…nd) have to be determined. This vector represents a joint state and its first elements express risk factor class for each risk factor r=1,…nrd, while its the last elements express disease state (with the disease=1, without the disease=0) for each disease d=A,B,..nd.

In principle, the state prevalences could be based on empirical data. In practice, the available data will be less detailed. Therefore, the model also enables to approximate the joint prevalence rates from total disease prevalence and total risk factor prevalence rates in combination with relative risks. If more detailed input data are available, these may be used instead.

The initial class prevalence rates of the CDM2005 joint version were calculated in successive steps. We illustrate the results with an example on two risk factors with two classes each and two diseases.

(1) We calculated the prevalence rates for all joint risk factor classes. In case of two risk factors the class prevalence rates generate a two-dimensional table.

Table 1: Class prevalence and marginal prevalence in case of two risk factors

		risk factor 2		marginal
		class 1	class 2	
risk factor 1	class 1	p_{11}	p_{12}	p^1_1
	class2	p_{21}	p_{22}	p^1_2
marginal		p^2_1	p^2_2	

with:

p_{ij} prevalence rate for class i for risk factor 1 and class j for risk factor 2

p^r_i class i prevalence rate for risk factor r

By applying assumption 3 of independent risk factor distributions we find:

$$p_{ij} = p^1_i \, p^2_j$$

That means, the joint risk factor class prevalence rate is the product of the class prevalence rates. This first step can be written in general terms of any number of risk factors selected:

$$p_0(\underline{r}) = \Pi_r \, p^r_i$$

with:

$\underline{r}$ vector of classes i for all risk factors r distinguished

p^r_i class i prevalence rate for risk factor r

$p_0(\underline{r})$ joint prevalence rate of state $\underline{r}$

(2) We multiplied these prevalence rates with the disease state prevalence rates applying the assumption of independent disease distributions conditional on the risk

factors included. In case of two diseases each joint risk factor class (i,j) generates a two-dimensional table on all joint disease states.

Table 2 : Prevalence for joint states in case of two diseases

risk factor classes i, j		disease 2		
		with	without	
disease 1	with	$p_{00}(i,j)$	$p_{01}(i,j)$	
	without	$p_{10}(i,j)$	$p_{11}(i,j)$	
				$p_{ij} = p^1_i \, p^2_j$

with:

$p_{AB}(i,j)$ prevalence rate of state A for disease 1 and state B for disease 2 conditional on class i of risk factor 1 and class j of risk factor 2

$p_{ij} = p^1_i \, p^2_j$ joint prevalence rate of class i of risk factor 1 and class j of risk factor 2.

For the general case of nd diseases, we calculated the joint disease state prevalence rates by successively calculating the prevalence rates for each disease separately. That is, for disease d, the joint prevalence rate has to be multiplied with:

if the state of disease d is 0 (disease-free): $(1 - \Pi_r \, RM^r_{d,i} \, p_d)$

if state of disease d is 1 (with disease) then: $\Pi_r \, RM^r_{d,i} \, p_d$

with:

p_d disease d prevalence rate

$RM^r_{d,i} = RR^r_{d,i} / \Sigma_i \, RR^r_{d,i} \, p^r_i$ risk multiplier for disease d for class i of risk risk factor r

For the general case of nd diseases, hence results:

$p_0(\underline{r}, \underline{d}) = \Pi_r \, p^r_i * \Pi_{d,\, state=1}(\Pi_r \, RM^r_{d,i} \, p_d) * \Pi_{d,\, state=0}(1-\Pi_r \, RM^r_{d,i} \, p_d)$

with :

$\underline{d}$ vector of states for all diseases d distinguished

$p_0(\underline{r}, \underline{d})$ joint prevalence rate of state $(\underline{r}, \underline{d})$

(3) We adjusted these prevalence rates for co-morbidity for each pair of causally related diseases, conditional on all risk factors. That means, if two diseases are causally related, then the joint prevalence rates of having both diseases simultaneously or being disease-free are larger than the expected ones, i.e. assuming independent diseases. As a result, the prevalence rates of having one of both diseases only are smaller than the expected ones. These formulas are rather complex and not shown here.[16]

(4) Since the latter adjustment is made on all co-morbidity states aggregated over all other diseases and risk factors and not on each state separately, it is an approximation. Therefore we re-scaled the prevalence rates such that the sum over all disease states and risk factor classes equals 1.

3.3.1.7 Incidence rates for all model states

Using these prevalence rates we calculated the incidence rates for each model state. The general idea is that all state-specific incidence rates are proportional to an unknown baseline incidence rate such that the weighted sum of the class-specific incidence rates, with weights equal to the initial state prevalence rates, is equal to the known overall incidence rate. That is:

$inc_d(\underline{r},\underline{d}) = inc_0 * \Pi_r RR^r_{d,i} \Pi_D RR_{d,D}$

and inc_0 follows from

$inc_d = \Sigma_{\underline{r},\underline{d}} inc_d(\underline{r},\underline{d}) * p_0(\underline{r}, \underline{d}) = \Sigma_{\underline{r},\underline{d}} inc_0 * \Pi_r RR^r_{d,i} \Pi_D RR_{d,D} * p_0(\underline{r}, \underline{d})$

so that :

$inc_d(\underline{r},\underline{d}) = \{ \Pi_r RR^r_{d,i} \Pi_D RR_{d,D} \ / \Sigma_{\underline{r},\underline{d}} p_0(\underline{r}, \underline{d}) \Pi_r RR^r_{d,i} \Pi_D RR_{d,D} \} * inc_d$

with:

inc_0	baseline incidence rate
$\underline{r}$	vector of classes i for each risk factor r included
$\underline{d}$	vector of states (1/0) for each disease d included
$(\underline{r},\underline{d})$	joint model state
$inc_d(\underline{r},\underline{d})$	disease d incidence rate for joint model state $(\underline{r},\underline{d})$
inc_d	disease d incidence rate
$p_0(\underline{r}, \underline{d})$	prevalence of joint model state $(\underline{r},\underline{d})$
$RR^r_{d,i}$	relative disease d incidence risk for class i of risk factor r
$RR_{d,D}$	relative disease d incidence risk for disease D conditional on all risk factors. Diseases D are all intermediate diseases for disease d.

For diabetes, for some risk factors the disease incidence rate equals 1, while it is known that the prevalence distribution in diabetes patients differs from a general population of the same age and gender structure. This is the case for cholesterol and blood pressure (see 3.2.3). This may be due to risk factor clustering, that is, the assumption of independent risk factors was not met. A solution to cope with this problem without too much effect on the overall structure of the model is to redistribute incidence of diabetes over the risk factor classes. Such a redistribution is also needed for the division of diabetes incidence over HbA1c classes.

3.3.1.8 Summary of model initialization

After the initialization steps have been taken, the following variables are available in the model:

$inc_d(\underline{r},\underline{d})$	incidence rates for all joint model states $(\underline{r},\underline{d})$
$p_0(\underline{r},\underline{d})$	prevalence rates for all joint model states
$m_{oc,0}$	the mortality rate for other causes of death, in a baseline state of no increased risk
$RR^r_{i,oc}$	relative risks for other causes of death, for classes i of risk factors r.
$RR^r_{d,i}$	relative disease d incidence risk for class i of risk factor r

$RR_{d,D}$	relative disease d incidence risk for disease D conditional on all risk factors. Diseases D are all intermediate diseases for disease d.
em_{d0}	Adjusted excess mortality rates for all diseases d
cf_d	the disease d case fatality rate, i.e. the 1-month mortality rate after disease onset, for d=AMI and d=CVA.
rem_d	disease d remission rates. The rate of remission equals zero for all diseases involved in the diabetes module.
$\lambda^r_{ij}(\underline{r},\underline{d})$	transition rates between risk factor classes, for each risk factor r and all classes i,j.

The values for these variables are partly specific to the choices of the risk factors and diseases to be included in the simulations and therefore they were calculated in the initialization for each simulation. Based on these variables the simulations can be run, calculating the changes of the joint prevalence rates over time.

3.3.2 Model simulation steps

In this paragraph we describe the model simulation part, i.e. the transitions over time between all model states. That means, between risk factor classes and disease states simultaneously. The mathematical model was set up in continuous time. Differential equations are used to describe the changes of the prevalence rates for any model state during small time steps. The time step is chosen sufficiently small, so that the probability of two events is very small compared to the probability of only one event. As a result, these differential equations describe the probability of any (one) event during the time step.

We describe the following aspects of the model simulation part:

- an example differential equation for the case of a model with two risk factors and two diseases
- the calculation of the transition rate matrix from the given transition rates for each event
- the general form of the CDM2005 differential equations

3.3.2.1 Example differential equations for two risk factors two diseases model

This example shows for the simple case of two independent diseases, how the prevalence rates change over time. We applied all assumptions described in section 3.1 to derive the differential equation that describes the change of the prevalence rate for any model state over a small time step[39]:

$p(i,j,d,e)_{t+\Delta t} - p(i,j,d,e)_t =$

disease-related adjusted excess mortality

$$(\text{if } d=1) - em_{d0}\, p(i,j,d,e)_t\, \Delta t +$$
$$(\text{if } e=1) - em_{e0}\, p(i,j,d,e)_t\, \Delta t +$$

other causes mortality

$$-RR^{(1)}_{oc,i}\, RR^{(2)}_{oc,j}\, m_{oc0}\, p(i,j,d,e)_t\, \Delta t +$$

disease incidence

$$(\text{if } d=0) \qquad - RR^1_{d,i}\, RR^2_{d,j}\, inc_{d0}\, p(i,j,0,e)_t\, \Delta t +$$

(if d=1) $+ RR^{1}_{d,i}\, RR^{2}_{d,j}(t)\, inc_{d0}\, p(i,j,0,e)_t\, \Delta t$
(if e=0) $- RR^{1}_{e,i}\, RR^{2}_{e,j}\, inc_{e0}\, p(i,j,d,0)_t\, \Delta t +$
(if e=1) $+ RR^{1}_{e,i}\, RR^{2}_{e,j}\, inc_{e0}\, p(i,j,d,0)_t\, \Delta t +$

risk factor class transition (outflow and inflow)

$- \Sigma_{u \neq i}\, \lambda^{(1)}_{iu}(t)\, p(i,j,d,e)_t\, \Delta t +$
$- \Sigma_{j \neq v}\, \lambda^{(2)}_{jv}(t)\,)\, p(i,j,d,e)_t\, \Delta t +$
$\Sigma_{u \neq i}\, \lambda^{(1)}_{ui}(t)\, p(i,j,d,e)_t\, \Delta t +$
$\Sigma_{v \neq j}\, \lambda^{(2)}_{vj}(t)\, p(i,j,d,e)_t\, \Delta t +$

$o(\Delta t)$

with:

t, Δt	time point and small time step respectively
i,j	classes (indexes) of risk factors 1 and 2 respectively
d,e	states (indexes) of diseases 1 and 2 respectively
$p(i,j,d,e)_t$	prevalence rate for model state (i,j,d,e) on time t
em_{d0}	adjusted disease d excess mortality rate
m_{oc0}	baseline mortality rate for other causes of death
inc_{d0}	baseline disease d incidence rate
$RR^{r}_{oc,i}$	other causes mortality relative risk for class i of risk factor r
$\lambda^{(r)}_{iu}(t)$	transition rate from class i to class u of risk factor r on time t
$o(\Delta t)$	denotes terms that are of secondary order or smaller in t

The transition rates between risk factor classes $\lambda^{(r)}_{iu}(t)$ were made time dependent to allow for the evaluation of scenarios that result in different transition rates for certain periods, for example, an increase of the smoking cessation rate during a 10 year time period. They can also be made dependent on the specific joint model state to model interventions in specific diseases, but this is not shown here.

3.3.2.2 General form of the CDM2005 differential equations

To describe the changes of the prevalence rates for any model state in a general model with nd diseases and nrd risk factors, all transition rates were combined into a nz*nz matrix, trans(t). Here nz denotes the number of different classes possible, which equals the multiplication over all risk factors and all diseases of the number of classes for risk factors (ncr), and the number of states for diseases (2) respectively: $nz = \Pi_r\, ncr(r) * 2^{nd}$. The details for the computation of the matrix trans are given in a background report.[16] Using this state transition rate matrix the differential equation that describes the changes of the model state prevalence rates in continuous time becomes:

$$p(z|t+\Delta t) = p(z|t) + p(z|t)\ .\ trans(t)\ \Delta t$$

with:

p(z|t)	the vector that describes the model state prevalence rates on time t
Δt	small time step

z=(r,d) the joint model states
trans(t) the matrix with transition rates between all classes

In the Mathematica model implementation, the model is run in one year time steps. A discretization procedure is used to transform the transition rates in continuous time to transition probabilities in one year time steps.[39]

3.3.2.3 Modeling disease duration

The CDM is a Markov type model. This implies that transitions were assumed independent of the residence time in model states. For example, the diabetes excess mortality rate (specific to age, gender and risk factor classes) does not depend on the duration of diabetes.
For a number of diabetes complications, it would be more realistic to have duration as an additional explanatory variable. However, a problem of correlation with age exists, so that a reliable estimation of both age and duration as explanatory variables based on empirical data is often difficult.
In CDM2005 joint version, diabetes duration is not yet included as an explanatory variable. Two possible ways to introduce duration into the model exist. Either additional states (for instance, first year of diabetes, less than 5 years diabetes, more than 5 years diabetes) may be introduced. Second, transition rates could be made dependent on the average residence time in the state. The second method is preferred, because in the first it is somewhat arbitrary how many states should be distinguished and the number of states may grow too large. The general method to compute average residence time in model states has been included in the model[41] The model is set up such that disease duration can be obtained as an outcome variable.

3.3.2.4 Modeling diabetes severity

Apart from its macro vascular complications, an important indicator of Diabetes severity and use of health care is the long term level of glycemic control as measured by HbA1c. In the current model version, diabetes is not divided into classes for HbA1c level. This will be included in a second model version, to enable the evaluation of interventions targeting at improved glycemic control.
It will involve the introduction of three Diabetes glycemic control stages, with cut-off points as reported by Baan et al..[29] Transition rates between these stages as well as the initial distribution of diabetes prevalence over the stages and the distribution of diabetes incidence have to be added to the model. Furthermore, it has to be considered whether complications and diabetes excess mortality depend on the level of glycemic control.

3.3.3 Effects of treatment.

The model was set up such that it is possible to analyze the effects of different types of interventions. The model as described above allows analyzing:

- interventions that affect risk factor prevalence in the general population and hence the incidence of DM. (primary prevention)
- interventions that affect risk factor prevalence in the DM population and hence the incidence of DM cardiovascular complications (tertiary prevention)

In a later stage, HbA1c will be included in the model and micro vascular complications will be linked to the HbA1c levels and included in the model. This will then allow analyzing:

- interventions that affect the level of glycemic control (HbA1c) and hence the prognosis of Dm over HbA1c classes
- interventions that affect the diagnosis of diabetes, resulting in diagnosis at earlier moments in time, so that the average disease duration will change

4. Budget allocation in a Chronic Disease Model: methodological issues

4.1 Background and state of the art

The goal of the project was to study budget allocation models as a tool for decision makers that have to allocate limited health care resources over a set of health care programs using applied cases to focus on relevant methodological issues. If more and more cost-effectiveness results become available, it is important to know how to combine the knowledge of the efficiency of different interventions. The explicit formulation of budget allocation models may be a necessary step to help decision makers combine the information from different cost-effectiveness analyses.

Budget allocation models formulate the allocation of budgets to healthcare interventions as the optimization of an objective, for example total quality of life in the population, subject to one or more constraints, for example the constraint that total costs should remain within a given budget, or the constraint that individuals should receive a certain minimum quality of care.

Budget allocation can be applied on any level of aggregation, with the scope of the optimization problem ranging from very broad (e.g. the question of the optimal level of health care budgets compared to for instance education budgets), to very specific (e.g. the optimal allocation of donor kidneys to different patient groups). At low levels of aggregation, budget allocation comes close to cost effectiveness analysis and the distinction between the two is artificial. However, many straightforward cost effectiveness analyses, especially if based on individual data, do not mention total budget impacts or total health effects and therefore can not be used for questions of budget allocation. They report for instance costs and health gains per patient and incremental cost effectiveness ratios. More complex cost effectiveness analyses, for instance, many cost effectiveness analyses of screening programs, usually do mention outcomes on a population level and contain elements of budget allocation. Evaluations of screening programs may for instance address the question of the most efficient choice of target group and screening interval.[42]
Conceptually, a ranking may be made from specific to broad scopes as pictured in Figure 12 below.

It may be expected that the broader the scope of the budget allocation analysis, the more methodological and data problems will arise. In the current report, the focus will be on intermediate levels of aggregation as marked by the grey area in Figure 12. This allows us to illustrate our budget allocation problems with applications based on empirical data. Furthermore, we will only consider chronic diseases. The methodological issues at stake differ depending on the focus (specific diseases, or specific types of treatment, or specific patient groups). In the case study of Diabetes for instance, we concentrate on different types of interventions for different target groups, but all aiming at a single disease. In another case study we intend to concentrate on a single type of intervention (primary prevention), for different target

groups and different chronic diseases. Each limitation will eliminate specific methodological problems, so that in each case study we will be able to focus on a small set of relevant methodological problems.

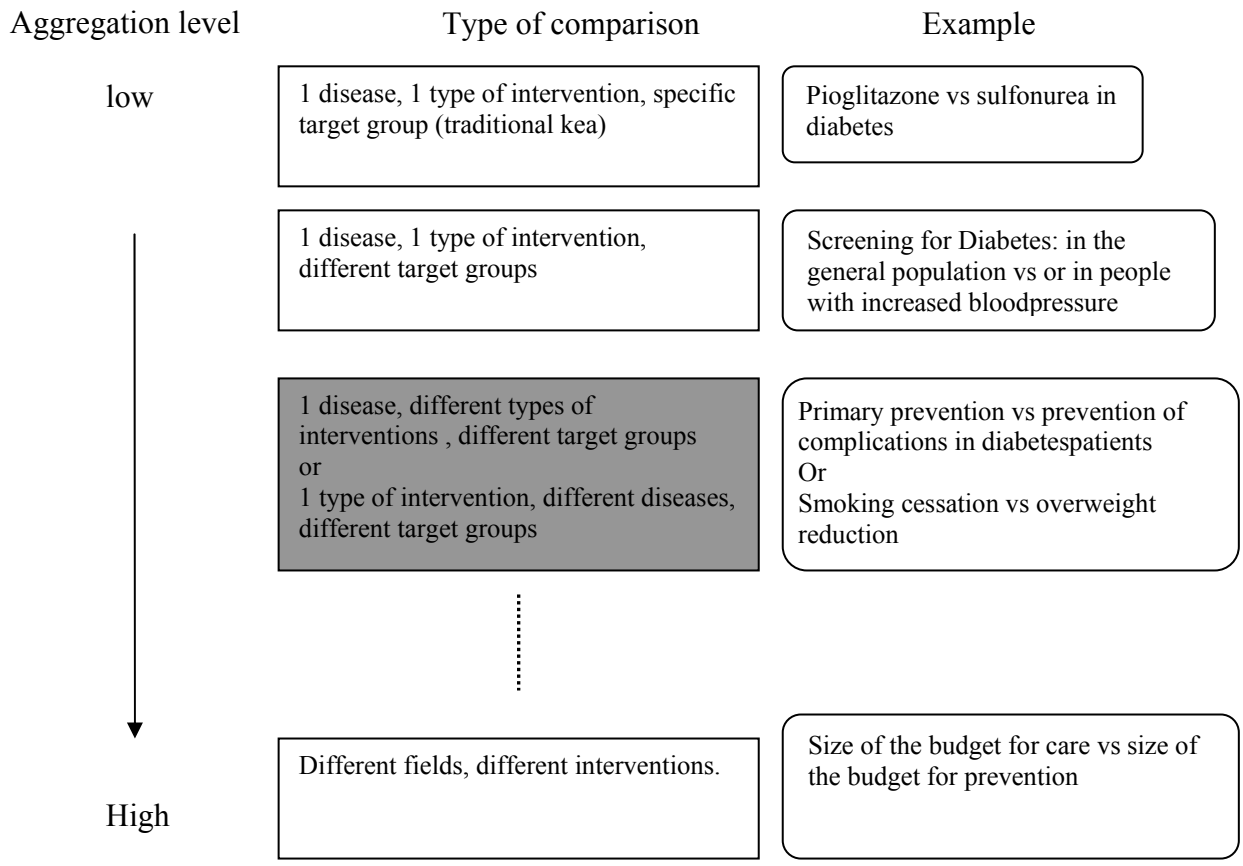


Figure 12: Scope of budget allocation problems

A literature search was performed to identify applications of budget allocation in the literature. Keywords used were resource allocation, model*, optimis(z)e, and costs* in different combinations.

We classified papers according to their scope:

- Broad. This refers to articles considering the allocation of different types of interventions, for different diseases and different target groups. Sometimes the budget at stake is the entire health care budget. Most of these applications apply to developing countries.
- Narrow. This refers to applications considering the allocation of a single type of interventions (usually prevention), to a single disease or disease category, for instance cardiovascular disease.
- Intermediate. This refers to applications with either a single type of interventions, or a single disease, or a single target group, and is the focus of our research.

A list of papers with their scope is presented in Table 3 below. Please note that this list is incomplete for the narrow applications, since our searches aimed to identify papers with intermediate and broad scopes.

Table 3: Overview of applied budget allocation models in the literature

Authors-publication year, country	Scope	Type of interventions	Disease(s)	Target groups
Murray et al. (1994), hypothetical developing country[43]	Broad	Infrastructure, primary prevention, screening, treatment	various	Differ per intervention
Granata et al. (1998), USA[44]	Broad	Primary prevention, screening, diagnosis, treatment	various	Differ per intervention
Flessa (2000), Tanzania[45]	broad	Primary prevention, Treatment	various	Differ per intervention
Feldstein, et al. (1973), Korea[46]	Intermediate	Primary prevention Treatment	Tuberculosis	4 age groups, and rural vs urban.
Barnum et al. (1980), Colombia[47]	Intermediate	Infrastructure, Primary prevention, Screening, Treatment,	various	Children: neonatal, infants and toddlers.
Cromwell et al. (1998), Australia[48]	Intermediate	Acute inpatient service	various	Differ per intervention
Niessen et al. (2005), NL[12]	Narrow/Intermediate	Prevention of complications, Treatment	Stroke	Patients in three disability states
Hutubessy et al.(2004), NL[11]	Narrow/Intermediate	Prevention of complications, Treatment	Diabetes	Patients in different disease stages
Earnshaw, et al. (2002), USA[15]	Narrow	Prevention of macro vascular complications	Diabetes	10 year age categories
Lindholm et al. (1999), Sweden[49]	Narrow	Primary prevention	Cardiovascular disease	groups acc to regio, age, gender and risk factor prevalence.
Marshall et al. (2002), UK[50]	Narrow	Primary prevention	Cardiovascular disease	Groups acc to risk score.
Murray et al. (2003), different large regions (e.g. Western Europe)[51]	Narrow	Primary Prevention	Cardiovascular disease	Population and groups acc to risk score
Richter et al. (1999), USA[52]	Narrow	Primary prevention	HIV	Injecting drug users vs. non, low or high risk
Zaric et al. (2001), USA[53]	Narrow	Primary Prevention	HIV	IDU, and subgroups vs. non-IDU

It may be concluded that the number of recent applications with an intermediate or broad scope is quite limited and has in the past been mostly focused on developing countries, for which of course, budget allocation questions are very relevant. For chronic diseases the number of budget allocation studies is not very large and almost entirely concerns primary prevention. An interesting exemption are the studies by the WHO[3], among which are those by Hutubessy et al. on diabetes.[11] However, even these studies do not really compare primary prevention and prevention of complications in the same model, as we intend to do.

Another interesting characteristic of published studies was that the effect of different overall budgets on the optimal solution was studied. It is one of the advantages of budget allocation models that this can be done relatively easy. All studies also analyze variants with additional constraints on the amount of money spent on subsets of programs. Barnum[47] analyses various resource and capacity constraints, and Murray et al.[43] has variants with and without capacity constraints. Furthermore, due to the static nature of most models, only once and for all solutions are found and time constraints on budgets cannot be dealt with.

In the following subsection, the budget allocation problem will be formalized as a mathematical programming problem. Such a formalization forces to be explicit about objectives, decision variables and constraints and enables to be precise in the formulation of the questions and methodological problems to be addressed. Then, we will discuss what adjustments are needed to realize a budget allocation model linked to the RIVM chronic disease model (CDM). Subsection 4.4 will introduce the methodological issues to be addressed in empsirical applications of budget allocation for chronic diseases at the intermediate level. Subsection 4.5 then shortly discusses the question which interventions are interesting from the point of view of budget allocation for the case study of diabetes.

4.2 Formalization of the budget allocation problem

The explicitly formulation of a budget allocation model means that the objectives and constraints of the decision maker have to be written down. The most straightforward way to do this formally is in the form of a so called 'mathematical programming problem'. This term covers several types of optimization problems, with linear programming problems being a well-known example. The standard model from Weinstein for instance, is a linear programming problem.[1]
An advantage of the explicit formalization in a mathematical programming problem is that the objectives, choices and constraints at stake are made explicit. In the literature, several variants of the standard model have been analyzed, including analyses on mutually exclusive interventions [54], on variable returns to scale[55],[56] indivisibility [57] and of the effect of uncertainty. [58-60]

In a mathematical programming or optimization problem, one seeks to minimize or maximize a real function of real or integer variables, subject to constraints on the variables.[61] Formalization of budget allocation problems as a mathematical programming problem implies therefore that the objectives and constraints have to be sorted out and that one objective has to be formulated in the form of a function that relates a single value to each possible combination of choices for the decision variables, and a set of constraints, which also depend upon the choices for the decision variables.

The mathematical programming problem can be formally denoted as follows:
Max f(d,v,w,t)
s.t. G(d,v,w,t)<=0

With :
t denotes time
$d=(d_i(t))$ is a vector of i=1,..n decision variables that for some problems may be chosen to depend on time (for instance, the number of people with diabetes which receive intensive treatment to reduce their dyslipidemia),
$v=(v_j(t))$ is a vector of j=1,..m variables which describe intermediate factors influencing the objective. These are variables which are influenced by the decision variables and in turn will affect the value of the objective function. (for instance, the number of people with diabetes and cardiovascular complications)
$w=(w_j(t))$ is a vector of j=1,..m variables which describe external factors influencing the objective (for instance, the number of births)

f(d,v,w,t) is the objective function, (for instance, the net present value of the number of life years lived by the population considered).
G(d,v,w,t)<=c denotes constraints on the choice of the decision variables in a very general form. G may be a single or a set of functions and defines the set from which the decision variables have to be chosen.
Simple budgetary constraints for instance may be modeled by using :
G(d,v,w,t)=$\sum_t$(((1/(1+r)^t)*d(t)*c(t))-B, where c(t) are the costs per patient at time t, d(t) the number of patients treated, B the total budget and the factor (1/(1+r)^t is used to compute the net present value of the total costs, and r is the rate of discount.

The mathematical programming formulation is flexible and allows for many variants of the standard model. For infectious diseases, variants have been analyzed by Brandeau and coauthors.[62] It is our aim to analyze variants applicable to chronic diseases as they were modeled in the RIVM Chronic Disease Model.

The standard budget allocation model as written down for instance by Weinstein and co-authors[1] is static and has maximization of the sum of health effects as its optimality criterion, under the constraint that the sum of program costs remains within a given total budget and, for the case of mutually exclusive interventions, may be written as a mathematical programming problem as follows.[54 60]

$$\underset{d_i^j}{Max} \quad \sum_{j=1}^{J} \sum_{i=n^{k(j)-1}}^{n^{k(j)}} d_i^j q_i^j \; p^j$$

subject to

$$b \geq 0$$

with $b = b(0) - \sum_{j=1}^{J} \sum_{i=n^{k(j)-1}}^{n^{k(j)}} d_i^j \; c_i^j \; p^j$ and p^j, *b(0)* given.

With:

j	Index for disease group, j=1,…J.
i	Index for programs, i=1,…n programs
$p^j(t)$	Number of patients in disease group j
k(j)	Index for the set of mutually exclusive programs for treatment of disease group j, i.e. J sets k(1), …k(J).
$n^{k(j)}$	Value of i for the last program that belongs to set k(j) and is meant for treatment of patients in disease group j.
$d_i^j(t)$	Fraction of patients in disease group j that receive treatment i
q_i^j	Effects of treatment i for patients in stage j
b	remaining budget; b(0) total available budget
c_i^j	Costs of treatment i in disease group j

A variant of this standard model explicitly shows the interdependencies between treatment choices and the number of patients in different stages of a disease is the formulation of a dynamic programming problem. Consider a chronic disease with a limited number of disease stages, for instance mild, moderate and severe disease. Assume that a Markov process can describe disease progression. Health effects are

expressed in effects on quality of life and mortality and measured as quality adjusted life years (QALYs). Then the following mathematic programming problem expresses the budget allocation problem of a decision maker that wants to choose treatments $d_i^j(t)$, the fractions of patients in each disease stage, j, that receive certain treatments, i, so as to maximize total health effects over the entire time horizon considered.

Notation :

j	Index for disease group/stage of disease, j=1,…J.
i	Index for programs, i=1,…n programs
t	Index for time periods, t=0,…T,
$p_j(t)$	Number of patients in disease group j at the end of period t
P(t)	P(t) =[pj(t)], a vector with the numbers of people in each disease stage
k(j)	Index for the set of mutually exclusive programs for treatment of disease group j, i.e. J sets k(1), …k(J).
$n_{k(j)}$	Value of i for the last program that belongs to set k(j) and is meant for treatment of patients in disease group j.
$d_i^j(t)$	Fraction of patients in disease group j that receive treatment i in period t
q_i^j	Direct effects of treatment i for patients in stage j
b(t)	Budget available at the end of period t
a_{jm}^i	Indirect effects of treatment i, that is, fraction of people in disease group j that progress to disease group m in the next period, if they receive treatment i.
A(t)	matrix with transition rates from each disease state into the other.
c_i^j	Costs of treatment i in disease group j
QY(t)	Number of quality adjusted life years, summed over all patients in all disease states, enjoyed up to period t inclusive.

$$\underset{d_i^j(t)}{Max} \quad QY(T) = \sum_{t=1}^{T} \frac{1}{(1+r)^t} \sum_{j=1}^{J} \sum_{i=n^{k(j)-1}}^{n^{k(j)}} d_i^j(t)\; q_i^j\; p^j(t-1)$$

subject to :

1
$$(a) \quad b(T) \geq 0$$
$$(b) \quad b(0) = bo$$

2. $0 \leq d_i^j(t) \leq 1$, for all j=1,...,J, for all i, for all t.
3. $\sum_{i=n^{k(j)-1}}^{n^{k(j)}} d_i^j(t) = 1$, for all j=1,...,J, for all t=1,…,T
4. $b(t) = b(t-1) - TC(t)/(1+r)^t$, t=1,…,T
5. $a_{jm}(t) = \sum_{i=n^{k(j)-1}}^{n^{k(j)}} d_i^j(t) a^i{}_{jm}$.
6. $P(t) = A(t)P(t-1)$, t=1,…,T; P(0) given
7. $QY(t) = QY(t-1) + TQ(t)/(1+r)^t$, t=1,…,T

with total treatment costs per year, TC(t), defined as the sum of costs over all disease states:

$$TC(t) = \sum_{j=1}^{J} \sum_{i=n^{k(j)-1}}^{n^{k(j)}} d_i^j(t) c_i^j\; p^j(t-1) \quad \text{, for all t.}$$

and the total number of QALYs enjoyed per year, TQ(t) can be found as the sum over all disease states of quality of life in each state:

$$TQ(t) = \sum_{j=1}^{J} \sum_{i=n^{k(j)-1}}^{n^{k(j)}} d_i^j(t) q_i^j \; p^j(t-1) \text{ , for all t.}$$

Here, constraint 1(a) represents the restriction that the net present value osf the total amount of money spent over the entire planning horizon should not surpass a given limit, or in other words, that at the end of the planning horizon a positive (or ideally zero) amount of money is left. The budget available at the start of the first period is given by b(0). Constraint 2 indicates that the fractions chosen should be between zero and one. Constraint 3 indicates that all patients in a disease state, j, must get some treatment.
Equation 4 defines the available budget at the end of year t as the budget at the end of the previous period minus total treatment costs in the current period. Equation 5 represents the assumption that the transition rates $a_{jm}(t)$, the elements in the matrix A, equal the weighted average of given, constant, treatment specific transition rates, a^i_{jm}. Equation 6 expresses that the distribution of patients at start is given by P(0), and for each time period follows from the application of transition rates to the distribution in the period before. Equation 7, finally, defines the accumulation of health effects: at the start of the first year, *QY(0)* is defined equal to zero, while at the end of the last year, *QY(T)* denotes the total number of quality adjusted life years enjoyed by all patients in all years, or total health effects.

This variant allows deriving some general notions about the direct and indirect effects of treatments and how these will be weighted in budget allocation decisions.

4.3 Budget allocation in the RIVM CDM.

This section addresses the question what adjustments to the RIVM Chronic Disease Model (CDM) are needed to enable optimization over costs and health effects. The RIVM Chronic Disease Model was developed for scenario analysis. It is structured as a simulation model that computes a number of outcomes (population numbers, the combined prevalences of diseases and risk factors, and mortality as well as secondary outcomes like life expectancy, quality of life and costs of diseases) for a range of years. Different scenarios can be formulated, for instance with different prevalence of risk factors, different transitions between risk factors, or different relative risks for complications to describe the effect of interventions. The model simulation will then be run and outcomes will be computed for each scenario respectively.

To enable budget allocation with the Chronic Disease Model several steps have to be taken:

- The first step is an extension of the current chronic disease model with a health economics module which computes outcomes in terms of life years gained, QALYs gained and health care costs.
- The second step is to include an objective function in terms of these health economic outcomes.
- The third step is to consider which parameters in the model may change as a result of interventions, that is, what are the decision variables that may be adjusted to optimize the objective function?

- The fourth step is to define the constraints that must be taken into account, when trying to optimize the objective function.
- The fifth step is the actual optimization, that is choosing the decision variables such that the objective is optimized, within the limitations defined by the constraints.

The first step has been addressed this year and will be documented in a separate report entitled 'Cost Effectiveness Analysis with the RIVM Chronic Disease Model'.[63] A health economics module was developed to integrate the cost-effectiveness calculations into the Mathematica implementation of the Chronic Disease Model. It computes total costs and total effects in terms of life years gained and QALYs for each intervention analyzed, as well as the comparison with a baseline scenario which results in incremental costs, incremental effects and cost effectiveness ratios. The module is flexible and allows for instance to vary the rate of discount, the time horizon and the approach to the inclusion of survivor costs and effects.

The second step involves the formulation of objective functions. In the project 'Budget allocation under uncertainty and time constraints', which was subsidized by ZONMW, we have compared several objective functions. A review of the literature on preferences and criteria of health care decision makers helped to formulate a number of objective functions, which have been discussed with an expert panel in a set of four semi structured interviews.[64] Based on these interviews, the following optimality criteria seemed relevant:

1 Maximize health effects (given budget constraints)

$\max \sum_{i=1}^{M} n_i * qaly_i$, with i an intervention, n_i the number of patients using intervention i, $qaly_i$ the average health gain per person with intervention i and M the total number of interventions considered.

It is assumed that the goal of the decision maker is to maximize total health effects. This is the basic premise on which the idea to calculate cost-effectiveness ratios is based. Health effects are defined as the increase in health (life expectancy or quality of life) as a result of a health care intervention, measured through the QALY (quality adjusted life year). Not only the number of QALYs per patient, but also the total number of QALYs from a program and hence the expected number of participants is important. Given the presence of a constraint on the total budget, the implication of this objective is that whenever new, efficient, health care programs become available, some currently implemented programs must be abandoned .

2. Maximize net present value

$\max \sum_{i=1}^{M} \gamma * n_i * qaly_i - \sum_{i=1}^{M} n_i * \mathrm{cost}_i$, with $cost_i$ as the average costs per person with intervention i, γ the fixed threshold ratio, and the other parameters as above.

Now the total value (i.e. gains minus costs) is maximized. A fixed willingness-to-pay value or 'exchange rate' is used to convert health effects into a monetary value. Note that this involves the essential assumption that health effects can indeed be converted into monetary values. The use of this objective is analogous to the decision rule to implement programs if their cost-effectiveness is below a fixed threshold value.[60] Because all health care programs with a ratio of costs to effects lower than the fixed

threshold ratio are implemented, the budget will increase whenever new, efficient, programs become available. The objective is also known under the name of 'net benefit approach', and is characterized by a flexible budget with fixed threshold ratios.

3. Maximize net present value with weighted health effects

$$\max \sum_{c=1}^{C} \gamma_c \sum_{i=1}^{M_c} n_{ic} * qaly_{ic} - \sum_{c=1}^{C} \sum_{i=1}^{Mc} n_{ic} * cost_{ic}$$

Per class of necessity c=1,…,C, a fixed threshold ratio γ_c is used, and n_{ic}, $qaly_{ic}$ and $cost_{ic}$ are the number of patients, effects and costs per person, for intervention i in class c. Per class M_c interventions are available.

Here it is assumed that the willingness-to-pay (i.e. the threshold ratio) is not a constant, but depends on the average severity of disease of the patient group under consideration. That is, for a patient group that is very ill the willingness-to-pay for a new treatment is higher than for a disease that only marginally affects quality-of-life/length-of-life. This concept was studied by Stolk et al, for the Health Care Insurance Board,[65 66] as a variant of the previous goal. Severity of disease is defined by relating the expected number of QALYs for patients with a certain disease to the expected number of QALYs of a healthy person. This ratio then determines the 'class of necessity' of the disease. Per class of necessity (c=1,…,C), a fixed threshold ratio γ_c is used, with a higher ratio if a health care program is intended for a more severe disease. As with goal 2, the budget is flexible, and will increase whenever efficient programs become available.

4. Decreasing marginal value with weighted health effects

$\max V(\sum_{c=1}^{C} \sum_{i=1}^{N} n_{ic} * w_c * qaly_i, B - \sum_{c=1}^{C} \sum_{i=1}^{N} n_{ic} * \cos t_{ic})$, where classes of necessity are defined as in 3.

As was remarked at 2, a fixed threshold ratio may lead to ever increasing budgets. Therefore, instead of using a fixed ratio, a more flexible approach might be used by looking at the total health effects already attained and the total budget spent. Such an approach can be formalized by defining a value function over health and money (budget remaining available for purposes competing with health care spending), and then maximizing value.
The value function is characterized by decreasing marginal value from both money and health. This means that an increase in total health leads to a smaller increase in value if the current amount of health is already high. Likewise, a decrease in total health care costs leads to a larger increase in value if the current costs are high. Thus, if new efficient programs become available, the threshold ratio gradually decreases. With this approach, a balance is sought between 1 and 2, i.e. the budget is no longer fixed but it will not increase indefinitely as more efficient programs become available. Furthermore, the concept of necessity is included by weighting QALYs. QALYs gained in a program aimed at a severe disease are given more weight than QALYs gained in a program aimed at a less severe disease.

These criteria were formulated without explicit reference to time or decision variables, hence for actual application they must be made more precise. The outcomes

that can be obtained in the health economics module of the CDM, enable to calculate $qaly_i$, n_i, $cost_{ic}$, for each year, for each intervention scenario evaluated, so that the value of these objective functions can be computed once values for the gamma, or necessity weights are known.

The third step will be further addressed in our future research for this project. An example is the introduction of intensive counseling for smoking cessation for diabetes patients. The decision variable then is the percentage of smoking diabetes patients which receives the intensive counseling, for each year included in the analysis.

The fourth step has, like the second step been addressed in our previous project 'Budget allocation under uncertainty and time constraints'.[64] Based on the results from this project, the following budget constraints will be considered.

1. Overall constraint
This is the most general type of budget constraint, and involves only a constraint on the total budget used in a certain planning period.

Formally: $\sum_{i=1}^{M} n_i * cost_i \leq B$

Again n_i is the number of patients using intervention i, M the total number of interventions, and cost_i the costs per person for intervention i. B is the total available budget.

2. Constraints on partial budgets for successive periods
An additional constraint maybe added if budgets are limited for each part of the planning period. For instance, a government may set up a general budget for a four-year period, with sub-budgets for each year. This kind of constraints is especially relevant when prevention programs have to be weighted against programs with an immediate effect on health.

Formally: $\sum_{i=1}^{M} n_i(t) * cost_i(t) \leq B(t)$ for $t=1,..T$

Here, t is the time period, with T the total planning period. Costs per person per intervention may differ per time period. B(t) is the budget available in time period t, and these budgets are fixed for all T time periods.

3. Upper and lower budget constraints
In contrast to the previous constraint here costs per period are bounded by both an upper limit and a lower limit. This is because in practice it may be important for the decision maker to spent most of the budget that was set aside for a certain period, thus making sure that the budget for the next periods is not decreased.

Formally: $b(t) \leq \sum_{i=1}^{M} n_i(t) * cost_i(t) \leq B(t)$ *for* $t=1,..T$

$b(t)$ and $B(t)$ are the minimum and maximum budget per time period.

4. Constraints for partial budgets for specific diseases or specific patient groups
This is similar to 2, but now with the total budget divided over patient groups instead of over time periods. For instance, sub-budgets may be defined for several broad disease area's, e.g. cardio-vascular, cancer, or vaccination.

Formally: $\sum_{i=1}^{M} n_i(c) * cost_i \leq B(c)$ *for c=1,..C*

C is the total number of specific groups, B(c) the partial budget for each group and $n_i(c)$ the number of patients in group c, using intervention i. Since not all M programs are relevant for all groups, $n_i(c)$ may be zero. We assume that the costs per person are the same for all patient groups.

5. Constraints on budget due to prior commitments
In practice, it is often not possible to freely (re)allocate the complete budget, as there may be prior commitments. This may for instance be the case for hospitals and nursing homes, where costs occur regardless of which programs is implemented. Likewise, budget may be committed to the care sector as a clear choice by society instead of being allocated according to some goal. In this way, some cost components of a health care program may be financed by these prior commitments, leaving only the remaining cost components to be provided for through budget allocation. We assume that it is even possible for all costs related to an intervention to be fixed through prior commitments.

Formally: $\sum_{i=1}^{M} n_i * mcost_i + V \leq B$

Here V is the total of prior commitments (which is thus not free to (re)allocate) and $mcost_i$ the part of the costs for intervention i that is not already covered in the prior commitments. This might take value zero, if all costs related to an intervention are fixed through prior commitments.

6. Constraint on partial budgets, coming from various sources.
This constraint is similar to 2 and 4. Different types of program costs may be financed from various sources (compartments), each with their own budget constraints. For instance, there may be separate budgets for pharmaceuticals, hospitals, home care etcetera, and an intervention may imply costs in each of these compartments. Such constraint will be relevant when a new program results in cost saving in one compartment and additional costs in another and thus shifts costs from one budget compartment to another.

Formally $\sum_{i=1}^{M} n_i * cost_{ji} \leq B_j$ *for j=1,...J*

Here j is the source and $cost_{ji}$ are the costs of program i that are paid for by source j, this may be zero. B_j is the budget of source j.

Some of these constraints reflect current practice, while others are interesting as a contrast. Their use in budget allocation exercises may show the effect of more flexible constraints.

Finally, the fifth step will be addressed in our further research for this project. A number of questions will have to be answered. One of these is whether we should optimize directly over the CDM, or it is better to use meta-modeling. In the latter case, first scenarios are run with CDM, then regression analysis is used to estimate a simplified relation between the outcomes and the decision variables, followed by optimization of this meta-model.

As a start, we evaluated what methods and software were used to compute optima in the literature and whether or not meta modeling was applied and if so, for what reason. The earliest applications were programmed in basic programming languages such as Fortran.[46 47] More recent applications sometimes used Excel solver for optimization.[44] This limits the complexity of the problems that can be addressed. Others applied own programs with more sophisticated optimization algorithms.[3 11 12] The more complex, nonlinear models on HIV prevention were solved in Mathematica.[52 53] Finally, GAMS and related software have been applied.[43], while some papers did not clearly report what software was used.[45 49]
A number of papers first computed outcomes for a list of scenarios, using a complex disease model, and then formulated the optimization problem as a choice over these scenarios, thus applying a simple form of meta-modeling. Others first simplified the disease model, using regression analysis to estimate relations between the decision variables and the outcomes, and then optimized using these relations. The remaining models were directly optimized, or the approach to optimization was not clearly reported.

4.4 Methodological issues

This section will shortly introduce a selection of the methodological issues to be addressed in real world applications of budget allocation at the intermediate level. Many more issues exist, but we selected the ones listed below based on their relevance and on the existing methodological literature. We wanted to address issues that had not been discussed extensively already (see page 42). It is not our intention to tackle the selected issues immediately. They will be the subject of research during the remainder of the project.

Time constraints on budgets
The standard model is static and therefore not fit for the analysis of constraints on the budget spent in certain periods. We plan to give a more dynamic formulation of the budget allocation problem. A first attempt for a formal description of such a dynamic budget allocation problem was given above.

Objective functions and different budgetary constraints
Optimality criteria other than maximization of total health effects are important. Above, we listed several objective functions that are potentially interesting, based on interviews with health care decision makers. It is our idea to run the optimization problems with each of these objectives to find out the differences in the optimal solutions found.

System limits
Every economic evaluation of a health care intervention is a partial analysis. It is neither possible nor useful to include all possible effects throughout the entire economy. The data requirements would be huge and the added precision small. The problem, as in applied cost-benefit analysis is to choose the effects that are included, so that only relevant effects are included and that no relevant effects are ignored.[67]

One of the methodological issues in costs-effectiveness analysis is whether or not to include indirect medical costs. Actual practice seems to be to exclude these costs in cost effectiveness analyses, mostly for pragmatic reasons. However, recent methodological work points into the direction of inclusion.[68 69] Especially for interventions in the area of prevention, for instance smoking cessation interventions, indirect medical costs may be substantial. Moreover, health gains may also depend on unrelated medical care. An interesting approach to this issue was recently set out by Nyman.[69] He argues that all costs that directly produce the utility measured in the denominator have to be included in the numerator of the cost effectiveness ratio. In the report [ref kea mo] a method is developed to consistently deal with direct and indirect health costs and effects within the health economic module of the RIVM Chronic Disease Model.[63]
Whether or not indirect medical costs should be taken into account is also important for budget allocation and is connected to the budget that should be optimally allocated. If one wants to optimally allocate the entire health budget it is no more than logical that indirect medical costs should be taken into account since they are part of the entire health budget. However, if the aim is to optimally allocate only a part of the health care budget (e.g. the prevention budget) then the question arises to what extent unrelated medical costs should be taken into account. An argument in favor of taking these into account would be that they also contribute to the health effects. However, strictly spoken they are not a part of the budget that should be optimally allocated.

4.5 Example interventions

Based on a review of existing cost effectiveness analyses, gaps in knowledge were identified.[14] The review focused on interventions aiming at macro vascular complications and at primary prevention. It was found that firm conclusions about cost-effectiveness seem possible for tight blood pressure control, with ratios ranging from cost savings to very low costs per LYG in six studies. Medication to reduce overweight and hyperglycemia in combination seems also cost-effective compared to conventional therapies, but this result was based on three studies. Finally, although the variation in interventions and evaluation methods was large, cost effectiveness ratios of medication against hyperglycemia were low in general except for new drugs. For the other interventions (primary prevention interventions, screening, and interventions in diabetes patients to reduce overweight, smoking, or dyslipidemia) more information is needed, either because few good quality studies were available or because methods and results differed too much between the different studies to enable a conclusion.

Therefore, interesting interventions to evaluate include all primary prevention interventions, as well as interventions for the prevention of complications in people with diabetes aiming at dyslipidemia, overweight and smoking.

5. Discussion and Conclusions

The current report described the building blocks to develop a budget allocation model for prevention in chronic diseases, including the prevention of complications in people with a disease. We choose to use a single model to evaluate all interventions. The model used was the RIVM Chronic Disease Model (CDM), and we illustrated the extensions needed to allow for budget allocation based on this model by the application in diabetes.
This report therefore first described a disease model for diabetes that allows evaluating both primary prevention and prevention of macro vascular complications. The report 'Modelling Chronic Diseases: the diabetes module. Justification of (new) input data in the mathematical RIVM Model' described the estimates of the input parameters from empirical data.[29] Together, these two reports document the modeling of diabetes and its complications in the RIVM Chronic Disease Model 2005 joint version.
Second, for the purpose of budget allocation, a health economics module was developed to integrate the cost-effectiveness calculations into the implementation of the Chronic Disease Model.[63] The module is flexible and allows for instance to vary the rate of discount, the time horizon and the approach to the inclusion of survivor costs and effects. It is also useful for other economic evaluations to be undertaken with the RIVM CDM.
Finally, the report introduced objective functions and budget constraints. To formulate objectives and constraints we could use our earlier work on budget allocation. They were derived from a theoretical model, but they have been discussed with health care decision makers.[64] The diabetes case study will enable to test their usefulness in an application.
Together with a choice on the decision variables and an approach to optimization, these elements form a budget allocation model.

From the WHO CHOICE project, it can be learned that it is very important to be as consistent as possible in the economic evaluation methods used for budget allocation. A good method for this is the use of a single epidemiological model to evaluate all interventions, like the CDM model in our case of diabetes.[3]
The RIVM Chronic Disease Model has a uniform structure for all modeled diseases, so that the work which has been done on diabetes easily applied to other modeled diseases as well. A similar approach has been advocated by the WHO[3], who apply their own model, POPmod. Compared to POPmod, the RIVM CDM is a bit more extensive, especially since it explicitly models risk factor prevalence in the population and links this to the incidence of diseases. Furthermore the CDM reflects the Dutch situation, while variants of POPmod may reflect the situation in various large regions in the world, for instance Western Europe. The CDM was intended primarily for evaluation of primary prevention interventions. Therefore we set out to extend the CDM to allow for the evaluation of interventions for the prevention of complications in people with diabetes.
A disadvantage of the evaluation of interventions in a single disease model for the purpose of budget allocation may be that the modeling work is very time consuming. However, gathering data on cost-effectiveness ratios from the literature, which would

be an alternative, is also time consuming, and the large problem of this approach is that cost-effectiveness ratios from foreign countries cannot usually be transferred to the national setting without a thorough consideration of the situation in each country and the methods used in the evaluation.[70] Furthermore, an approach without disease modeling limits the interventions that can be considered in the budget allocation exercise to those interventions for which good quality economic evaluations are already available.

The application in diabetes forced us to pay attention to the practical usefulness of theoretical concepts. For instance, we could quickly see if certain formal modeling required data that would not be available. The combination of this work with work on a report for the Ministry of Health enables to pay attention to both the theoretical and the empirical aspects. Our approach had the advantages of precision and the possibility to model interventions relatively straightforward. The disadvantage that relatively much input data were needed, was partly solved because we could start from the existing RIVM Chronic Disease Model.

The joint model that was developed for diabetes to keep track of the risk factor levels within patient populations is applicable to the other diseases modeled in the CDM as well. The CDM joint version allows comparing interventions in patients with primary prevention interventions in the general population. For example, for cardiovascular diseases, in the joint model, the smoking prevalences of patients are known and cessation interventions in patients may be evaluated. The joint modeling method that was developed to model diabetes is therefore useful for all diseases for which interventions on risk factors for patients are an interesting possibility. Of course, the data used will have to be validated with empirical estimates, like in the case of diabetes and this is quite a job.[29] For most diseases, a complete budget allocation model will also have to contain disease stages to enable the evaluation of treatment interventions, and this element is disease specific too. For diabetes, it is addressed by the modeling of different HbA1c classes.
In a second case, different interventions in primary prevention (smoking cessation and interventions to reduce overweight) will be compared, hence limiting the scope to a single type of intervention rather than to a single disease.

The modeling of diabetes for budget allocation proved to be a fruitful case study, and the joint model can in principle be applied to any modeled disease in the RIVM Chronic Disease Model. For all these diseases, the CDM2005 joint version allows to analyze interventions on risk factors in patients and compare them to primary prevention. Therefore, we conclude that the specific diabetes case study has a wider potential applicability. This also holds for the health economics module and the methodology on survivor costs and effects that will be applied for the economic evaluations along the CDM for the Ministry of Health. As regards the pure budget allocation elements, the objective functions, budget constraints and optimization approach, their discussion in the current report is too limited to allow for conclusions about their wider applicability. They will be the topic of our future research.

References

1. Weinstein MC, Zeckhauser R. Critical ratios and efficient allocation. J Public Economics 1973; 2:147-57.

2. Hoogenveen RT , de Hollander AEM, van Genugten MLL. The chronic disease modelling approach. Bilthoven: National Institute for Public Health and the Environment (RIVM) 1998.

3. Geneva: World Health Organisation, 2003.

4. Murray CJL, Lopez AD. The global burden of disease : a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020. Volume I. Cambridge, MA: Harvard School of Public Health, Published on behalf of the World Health Organization and the World Bank, 1996. (Global burden of disease and injury series.

5. Murray CJL, Lopez AD. Alternative projections of mortality and disability by cause1990-2020: Global Burden of Disease Study. Lancet 1997; 349:1498-504.

6. Baan CA, Feskens EJM. Disease burden of diabetes mellitus type II in the Netherlands: incidence, prevalence and mortality. Ned Tijdschr Geneeskd 2001; 145(35):1681-5.

7. Baan CA, Feskens EJM. Prevention of diabetes mellitus type 2. Ned Tijdschr Geneeskd 2001; 145(35):1677-80.

8. Seidell JC. Time trends in obesity: an epidemiological perspective. Horm Metab Res 1997; 29(4):155-8.

9. UK Prospective Diabetes Study (UKPDS) Group. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). UK Prospective Diabetes Study (UKPDS) Group. Lancet 1998; 352(9131):837-53.

10. Gaede P, Vedel P, Larsen N, Jensen GVH, Parving HH, Pedersen O. Multifactorial intervention and cardiovascular disease in patients with type 2 diabetes. N Engl J Med 2003; 348(5):383-93.

11. Hutubessy RC, Niessen LW, Dijkstra RF, Casparie TF, Rutten FF. Stochastic league tables: an application to diabetes interventions in the Netherlands. Health Econ 2004.

12. Niessen LW. Roads to health. Multi-state modelling of population health and resource use (PhD thesis). Amsterdam: Rozenberg Publishers, 2002.

13. Raikou M, McGuire A. The economics of screening and treatment in type 2 diabetes mellitus. Pharmacoeconomics 2003; 21(8):543-64.

14. Vijgen S, Hoogendoorn M, Baan C, de Wit G, Limburg W, Feenstra T. Cost-effectiveness of preventive interventions in Diabetes mellitus: a systematic literature review. Submitted 2005.

15. Earnshaw SR, Richter A, Sorensen SW et al. Optimal allocation of resources across four interventions for type 2 diabetes. Med Decis Making 2002; 22(5 Suppl):S80-S91.

16. Hoogenveen R . Chronic disease modelling. Implementation of the RIVM Chronic Disease Model 2005. Bilthoven: National Institute for Public Health and the Environment (RIVM), 2005; CDM technical report no 4.

17. Clarke PM, Gray AM, Briggs A et al. A model to estimate the lifetime health outcomes of patients with type 2 diabetes: the United Kingdom Prospective Diabetes Study (UKPDS) Outcomes Model (UKPDS no. 68). Diabetologia 2004; 47(10):1747-59.

18. Stevens RJ, Kothari V, Adler AI, Stratton IM, United Kingdom Prospective Diabetes Study (UKPDS) Group. The UKPDS risk engine: a model for the risk of coronary heart disease in Type II diabetes (UKPDS 56). Clin Sci (Lond) 2001; 101(6):671-9 + 679 + 681-682 + 217-219.

19. McEwan P, Williams JE, Griffiths JD et al. Evaluating the performance of the Framingham risk equations in a population with diabetes. Diabet Med 2004; 21(4):318-23.

20. DCCT Research Group. Diabetes Control and Complications Trial (DCCT). Update. DCCT Research Group. Diabetes Care 1990; 13(4):427-33.

21. Watkins P. The UKPDS. A model for gathering the evidence for the management of chronic diseases. UK Prospective Diabetes Study Group. J R Coll Physicians Lond 1998; 32(6):510-1.

22. Batchelder T, Barricks M. The Wisconsin Epidemiologic Study of Diabetic Retinopathy. Arch Ophthalmol 1995; 113(6):702-3.

23. Os Nv, Niessen LW, Koopmanschap MA, Lei Jvd. Detailed analysis of the societal costs of diabetes mellitus. Ned Tijdschr Geneeskd 2000; 144(18):842-6.

24. Palmer AJ, Roze S, Valentine WJ et al. Impact of changes in HbA1c, lipids and blood pressure on long-term outcomes in type 2 diabetes patients: an analysis using the CORE Diabetes Model. Curr Med Res Opin 2004; 20 Suppl 1:S53-8.

25. Palmer AJ, Roze S, Valentine WJ et al. Validation of the CORE Diabetes Model against epidemiological and clinical studies. Curr Med Res Opin 2004; 20 Suppl 1:S27-40.

26. Palmer AJ, Roze S, Valentine WJ et al. The CORE Diabetes Model: Projecting long-term clinical outcomes, costs and cost-effectiveness of interventions in diabetes mellitus (types 1 and 2) to support clinical and reimbursement decision-making. Curr Med Res Opin 2004; 20 Suppl 1:S5-26.

28. Kalbfleisch JD, Prentice RL. The Statistical Analysis of Failure Time Data . New York: Wiley, 2002.

29. Baan C, Bos G, Jacobs-vander Bruggen M Modelling Chronic Diseases: the diabetes module. Justification of (new) input data in the mathematical RIVM Model . Bilthoven: National Institute for Public Health and the Environment (RIVM), 2005; 260801001.

30. Statistics Netherlands . Statline [online database] [Web Page]. (Accessed 14 October 2003).

31. Stouthard M, Essink-Bot M, Bonsel G, Group. obotDDWD. Disability weights for diseases - a modified protocol and results for a Western European region. European Journal of Public Health 2000; (10):24-30.

32. Stouthard MEA, Essink-Bot ML, Bonsel GJ *et al*. Disability Weights for Diseases in The Netherlands. Department of Public Health. Erasmus University Rotterdam 1997.

33. Melse J, Kramers P. Berekeningen van de ziektelast in Nederland. Achtergronddocument bij VTV-1997 deel III, hoofdstuk 7 (Calculations of the burden of disease in the Netherlands. Background report for the Public Health Status and Forecast 1997 (report III, chapter 7)). Bilthoven: National Institute for Public Health and the Environment (RIVM), 1998; 431501028.

34. Melse JM, Essink-Bot ML, Kramers PGN, Hoeymans N, on behalf of the Dutch Burden of

Disease Group . A national burden of disease calculation: Dutch disability-adjusted life-years. Am J Public Health 2000; 90(8):1241-7.

35. Hoogenveen R . Chronic disease modelling. Relative risks and intermediate diseases. Bilthoven: National Institute for Public Health and the Environment (RIVM), 2005; CDM technical report no 2.

36. Hoogenveen RT, Gijsen R. Dutch DisMod for several types of cancer. Bilthoven: National Institute for Public Health and the Environment (RIVM) 2000.

37. Hoogenveen RT, Gijsen R., van Genugten MLL, Kommer GJ, Schouten JSAG, de Hollander AEM. Dutch DisMod. Constructing a set of consistent data for chronic disease modelling. Bilthoven: National Institute for Public Health and the Environment (RIVM) 2000.

38. Murray CJ, Lopez AD. Quantifying disability: data, methods and results. Bull World Health Organ 1994; 72(3):481-94.

39. Hoogenveen R . Chronic disease modelling. Adjusting excess mortality for comorbidity. Bilthoven: National Institute for Public Health and the Environment (RIVM), 2005; CDM technical report no 1.

40. Spasoff RA, McDowell IW. Estimating the combined effect of several disease precursors in health risk appraisal. Am J Prev Med 1987; 3(4):182-91.

41. Hoogenveen R , van Baal P, Bemelmans W. Smoking start, stop and relapse rates, analysis on retrospective data from StiVoRo. Bilthoven: National Institute for Public Health and the Environment (RIVM), 2005; CDM technical report no 3.

42. Gyrd-Hansen D. Cost-benefit analysis of mammography screening in Denmark based on discrete ranking data. Int J Technol Assess Health Care 2000; 16(3):811-21.

43. Murray CJL, Kreuser J, Whang W. Cost-effectiveness analysis and policy choices: investing in health systems. Bull World Health Organ 1994; 72 (4):663-74.

44. Granata AV, Hillman AL. Competing practice guidelines: using cost-effectiveness analysis to make optimal decisions. Ann Intern Med 1998; 128(1):56-63.

45. Flessa S. Where efficiency saves lives: a linear programme for the optimal allocation of health care resources in developing countries. Health Care Manag Sci 2000; 3(3):249-67.

46. Feldstein MS , Piot MA, Sundaresan TK. Resource allocation model for public health planning: A case study of tuberculosis control. Geneve: WHO , 1973; pp.109.

47. Barnum H, Barlow R, Fajardo L, Pradilla A. A resource allocation model for child survival. Oelgeschlager, Gum & Hair, Cambridge Dan. 1980.

48. Cromwell DA, Viney R, Halsall J, Hindle D. Linking measures of health gain to explicit priority setting by an area health service in Australia. Soc Sci Med 1998; 47(12):2067-74.

49. Lindholm L, Hallgren CG, Boman K, Markgren K, Weinehall L, Ogren JE. Cost-effectiveness analysis with defined budget: how to distribute resources for the prevention of cardiovascular disease? Health Policy 1999; 48(3):155-70.

50. Marshall T, Rouse A. Resource implications and health benefits of primary prevention strategies for cardiovascular disease in people aged 30 to 74: mathematical modelling study . BMJ 2002; 325(7357):197.

51. Murray CJL, Lauer JA, Hutubessy RCW et al. Effectiveness and costs of interventions to lower

systolic blood pressure and cholesterol: a global and regional analysis on reduction of cardiovascular-disease risk. Lancet 2003; 361(9359):717-25.

52. Richter A, Brandeau ML, Owens DK. An analysis of optimal resource allocation for prevention of infection with human immunodeficiency virus (HIV) in injection drug users and non-users. Med Decis Making 1999; 19(2):167-79.

53. Zaric GS, Brandeau ML. Optimal investment in a portfolio of HIV prevention programs. Med Decis Making 2001; 21(5):391-408.

54. Laska EM, Meisner M, Siegel C. The usefulness of average cost-effective ratios. Health Econ 1997; 6(5):497-504.

55. Elbasha EH, Messonnier ML. Cost-effectiveness analysis and health care resource allocation: decision rules under variable returns to scale. Health Econ 2004; 13(1):21-35.

56. Karlsson G, Johannesson M. Cost-effectiveness analysis and capital costs. Soc Sci Med 1998; 46(9):1183-91.

57. Birch S, Gafni A. Cost effectiveness/utility analyses. Do current decision rules lead us to where we want to be? J Health Econ 1992; 11(3):279-96.

58. Sendi P, Al MJ, Gafni A, Birch S. Optimizing a portfolio of health care programs in the presence of uncertainty and constrained resources. Social Science and Medicine 2003; 57, 2207-2215.

59. Sendi PP, Al MJ. Revisiting the decision rule of cost-effectiveness analysis under certainty and uncertainty. Social Science and Medicine 2003; 57, 969-974.

60. Al MJ, Feenstra TL, Hout BAv. Optimal allocation of resources over health care programs: dealing with decreasing marginal utility and uncertainty. Al MJ. Evaluating Health Care Technologies, PhD Thesis. 2001.

61. Mathematical Programming Society. What is Mathematical Programming? [Web Page]. Available at www.mathprog.org/mps_whatis.htm. (Accessed January 2005).

62. Brandeau ML, Zaric GS, Richter A. Resource allocation for control of infectious diseases in multiple independent populations: beyond cost-effectiveness analysis. J Health Econ 2003; 22(4):575-98.

63. van Baal P, Hoogenveen R, de Wit A, Feenstra T. Cost Effectiveness Analysis with the RIVM Chronic Disease Model. Bilthoven: National Institute for Public Health and the Environment (RIVM), 2005; 260706002.

64. Al MJ, Feenstra T, Brouwer WBF. Decision makers' views on health care objectives and budget constraints: results from a pilot study. Health Policy 2004; 70(1):33-48.

65. Toenders WGM. Breedte geneesmiddelenpakket (in Dutch). Amstelveen: College voor Zorgverzekeringen , 2001; publicatieno 01/54.

66. Toenders WGM . Vervolgonderzoek breedte geneesmiddelenpakket. Amstelveen: College voor Zorgverzekeringen, 2002.

67. Sugden R, Williams A. The principles of practical cost-benefit analysis. Oxford: Oxford University Press, 1978.

68. Meltzer D. Addressing uncertainty in medical cost-effectiveness analysis implications of expected utility maximization for methods to perform sensitivity analysis and the use of cost-effectiveness analysis to set priorities for medical research. J Health Econ 2001;

20(1):109-29.

69. Nyman JA, Hillson S, Stoner T, de Vries A. Do specialists order too many tests? The case of allergists and pediatric asthma. Ann Allergy Asthma Immunol 1997; 79(6):496-502.

70. Welte R, Feenstra T, Jager H, Leidl R. A decision chart for assessing and improving the transferability of economic evaluation results between countries. Pharmacoeconomics 2004; 22(13):857-76.

71. Boutayeb A, Twizell EH, Achouayb K, Chetouani A. A mathematical model for the burden of diabetes and its complications. Biomed Eng Online 2004; 3(1):20.

72. Kuo HS, Chang HJ, Chou P, Teng L, Chen TH. A Markov chain model to assess the efficacy of screening for non-insulin dependent diabetes mellitus (NIDDM). Int J Epidemiol 1999; 28(2):233-40.

73. Eastman RC, Javitt JC, Herman WH et al. Model of complications of NIDDM. I. Model construction and assumptions. Diabetes Care 1997; 20(5):725-34.

74. Eastman RC, Javitt JC, Herman WH et al. Model of complications of NIDDM. II. Analysis of the health benefits and cost-effectiveness of treating NIDDM with the goal of normoglycemia. Diabetes Care 1997; 20(5):735-44.

75. Stevens RJ, Kothari V, Adler AI, Stratton IM. The UKPDS risk engine: a model for the risk of coronary heart disease in Type II diabetes (UKPDS 56). Clin Sci (Lond) 2001; 101(6):671-9.

76. Brown JB, Russell A, Chan W, Pedula K, Aickin M. The global diabetes model: user friendly version 3.0. Diabetes Res Clin Pract 2000; 50 Suppl 3:S15-46.

77. Vijan S, Hofer TP, Hayward RA. Estimated benefits of glycemic control in microvascular complications in type 2 diabetes. Ann Intern Med 1997; 127(9):788-95.

78. Vijan S, Hofer TP, Hayward RA. Cost-utility analysis of screening intervals for diabetic retinopathy in patients with type 2 diabetes mellitus. JAMA 2000; 283(7):889-96.

79. Javitt JC, Aiello LP. Cost-effectiveness of detecting and treating diabetic retinopathy. Ann Intern Med 1996; 124(1 Pt 2):164-9.

80. Bagust A, Hopkinson PK, Maier W, Currie CJ. An economic model of the long-term health care burden of Type II diabetes. Diabetologia 2001; 44(12):2140-55.

81. Golan L, Birkmeyer JD, Welch HG. The cost-effectiveness of treating all patients with type 2 diabetes with angiotensin-converting enzyme inhibitors. Ann Intern Med 1999; 131(9):660-7.

82. Hauner H, Maxion-Bergemann S, Muller E, Schulz M, Huppertz E, Bergemann R. [Disease management program (DMP) diabetes mellitus: simulation of therapeutic results of different guidelines. A new diabetes mellitus model (DMM)]. Dtsch Med Wochenschr 2003; 128(21):1167-72.

83. Os Nv, Niessen LW, Bilo HJG, Casparie AF, Hout BAv. Diabetes nephropathy in the Netherlands: a cost effectiveness analysis of national clinical guidelines. Health Policy 2000; 51(3):135-47.

84. Davies R, Roderick P, Canning C, Brailsford S. The evaluation of screening policies for diabetic retinopathy using simulation. Diabet Med 2002; 19(9):762-70.

85. Murakami Y, Ohashi Y. Projected number of diabetic renal disease patients among insulin-

dependent diabetes mellitus children in Japan using a Markov model with probabilistic sensitivity analysis. Int J Epidemiol 2001; 30(5):1078-83.

86. Ragnarson Tennvall G, Apelqvist J. Prevention of diabetes-related foot ulcers and amputations: a cost-utility analysis based on Markov model simulations. Diabetologia 2001; 44(11):2077-87.

87. Ollendorf DA , Kotsanos JG, Wishner WJ et al. Potential economic benefits of lower-extremity amputation prevention strategies in diabetes. Diabetes Care 1998; 21(8):1240-5.

88. Gordois A, Scuffham P, Shearer A, Oglesby A, Tobian JA. The health care costs of diabetic peripheral neuropathy in the u.s. Diabetes Care 2003; 26(6):1790-5.

89. Gordois A, Scuffham P, Shearer A, Oglesby A. The health care costs of diabetic nephropathy in the United States and the United Kingdom. J Diabetes Complications 2004; 18(1):18-26.

90. Eddy DM, Schlessinger L. Archimedes: a trial-validated model of diabetes. Diabetes Care 2003; 26(11):3093-101.

91. Eddy DM, Schlessinger L. Validation of the archimedes diabetes model. Diabetes Care 2003; 26(11):3102-10.

Appendix 1 Models on complications of diabetes mellitus

This appendix summarizes the characteristics of diabetes models in the literature. It was based on a quick scan of the literature, not on a complete systematic review. Our aim was to identify and characterize models that included macro vascular complications and were modeling either unspecified diabetes or diabetes type 2.

The results were summarized in two tables, while for each model with at least some macro vascular complications included, a short summary of the model structure is given below. Table 4 presents the models found, categorized into three types of models: population models that do not specify complications, population or cohort models that specify complications, and so-called clinical models, which model the underlying clinical processes of diabetes in more detail than the other two types of models and lists the complications included for each model.

Table 4: Diabetes models, categorized into three types of models

Model	*Complications (source of parameter estimates)*	
	Macrovasc	*micorvasc*
Population models without specification of complications		
Morocco Model[71]	all complications, unspecified	
Taiwan model[72]	Symptomatic vs. asymptomatic (screened cohort data)	
Population or individual models with specified complications		
National Institutes of Health Model[73 74]	CVD (Framingham risk functions)	all 3 (WESDR / Rochester Study)
UKPDS Model[17 75]	CHD, CVA (Framingham risk functions, and UKPDS)	all 3
Global Diabetes Model (GDM)[76]	CVD (Framingham risk functions)	All 3 (as NIH model)
Vijan Model[77 78]		Retinopathy, nephropathy (NHANESIII, DCCT)
Prospective Population Health Event Tabulation Model (PROPHET)[79]		Retinopathy
Bagust Model[80]	CVD (Framingham risk functions)	all 3 (as NIH model)
Golan model[81]		Nephropathy (ACE inhibitor trials)
Diabetes mellitus Model[82]	AMI, CVA	all 3 (WESDR, DCCT)
Swiss Model[25 26]	AMI, CVA (Framingham and UKPDS)	all 3 (DCCT, UKPDS)
iMTA model[12 83]		all 3 (as in NIH)
Diabetic Retinopathy Screening Project[84]		Retinopathy
Japan model[85]		Nephropathy
Lund model[86]		Neuropathy
Lilly model[87]		Neuropathy
York model[88 89]		Nephropathy
Physiological models		
Archimedes[90 91]		

All 3: Retinopathy, Nephropathy & Neuropathy, DCCT: Diabetes Control & Complications Trial, WESDR: Wisconsin Epidemiologic Study of Diabetic Retinopathy, UKPDS United Kingdom Prospective Diabetes Study, NHANES(III) National Health and Nutrition Examination Survey

Population models without specification of complications

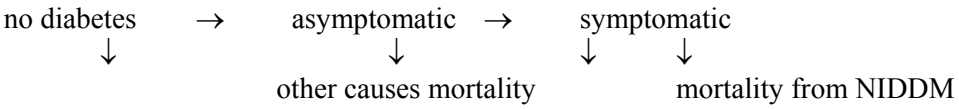
1. Morocco model

Deterministic time continuous Markov model (differential equations on class prevalence numbers), population cohorts. Based on Morocco population census & survey data.

Continuous-time differential equations of state probabilities:
$D'(t) = I(t) - (\lambda + \mu) D(t) + \gamma C(t)$
$C'(t) = \lambda D(t) - (\gamma + \mu + \nu + \delta) C(t)$

With D(t) # diabetics without complications,
C(t) # diabetics with complications; N(t) = D(t) + C(t)
I(t) incidence number
λ complications incidence rate (= β C / N, β=1)
μ diabetes mortality rate
γ complications cure rate
ν severe incurable complications incidence rate
δ complications excess mortality rate

2. Taiwan model
Deterministic time continuous Markov model (differential equations on class prevalence numbers), population cohorts. Based on cohort data from a population diabetes screening project (Pulio County, Taiwan).



Population or individual models with specified complications

1. National Institutes of Health Model (NIH, Eastman)
Stochastic discrete time (1 year) Markov model, individuals (micro simulation, Monte Carlo) Macro vascular complications modeled in a simple way. Transition probabilities were partly specified by duration diabetes (yr), partly specified by treatment. Based among other sources on (micro vascular) WESDR, Rochester Study, (CVD) Framingham risk function.

2. UKPDS model (on CHD complications)
Stochastic discrete time (1 year) Markov model, individuals, based on Framingham risk functions, and on UKPDS data.

Example of modeling of macro vascular complications:

NIDDM, no CHD duration t	$\rightarrow$p(t)	NIDDM, with CHD
	$\rightarrow$1-p(t)	NIDDM, no CHD, duration t+1

with :
p(t) 1-year first CHD event risk after t years since diabetes onset
$p(t) = 1 - \exp(- q \exp(d t)$
$\log(q) = \log(q_0) + \beta_1 (age-55) + \beta_2 gen + \beta_3 etn + \beta_4 cig + \beta_5 (HbA1c-6.72) + \beta_6 (SBP-13.57) + \beta_7 (ln(cholrat)-1.59)$
and age= age at diagnosis of diabetes, gen= gender, etn= ethnicity, cig= smoking status, SBP= systolic blood pressure, cholrat= lipid ratio, total/HDL cholesterol, t= years since diabetes onset with no CHD event, d= regression coefficient

3. Global Diabetes Model (GDM)
Discrete time (1 year) Markov-model, individuals, stochastic (Monte Carlo)
Structure differs from most other models

		1st hospital CVD event	→	next hospital event
	→			
	→	1st out-hospital CVD event	→	next out-hospital event
diabetes without complications				
	→	specific micro vascular event		

Model parts:
(1) yearly linear update of risk factors: new value = $b_0 + b_1$ * old value
(2a) first CVD event 1-year probabilities (based on 4-year Framingham CVD event risk function)
(2b) proportional distribution of first CVD event over manifestation form (source: Kaiser Permanente NW data), APE, CHF, AMI, PAD, CVA, for some forms divided into hospital or outpatient, percentages for singular event and including multiple events
(3) next CVD event rates (source: Kaiser Permanente NW data)
(3a) hospital events (#), using a Poison regression
(3b) out-hospital events (yes/no), using a logistic regression
(4) mortality probabilities

APE	angina
LEA	lower extremity amputation
PAD	peripheral artery disease

4. Diabetes Control and Complications Model (DCCT)
Based on NIH model (1) filled with DCCT data.

5. Bagust model
Deterministic discrete time (1 year) Markov model, deterministic, individuals. HbA1c classes based on therapy: H1: diet & exercise, H2: 1st line oral medication, H3: 2nd line therapy, H4: insulin-based therapy, HbA1c state transition probabilities. Macro vascular complications model based on Framingham risk functions

6. Diabetes mellitus Model (DMM)
Risk functions with time-dependent HbA1c levels, individuals (micro simulation, Monte Carlo), this is not a Markov-model. Data from WESDR, DCCT.

Diabetics (HbA1c, age) → AMI / CVA (non-fatal)

7. Swiss model (Palmer)
Stochastic time discrete (1 year) Markov-model, individuals (micro simulation, Monte Carlo)
Data from DCCT (micro vascular complications) and Framingham risk function (CVA, AMI).
Transition probabilities were partly specified for age, duration of diabetes, and treatment.

Example of modelling macrovascular complications:

No history of AMI	First AMI	Recurrent AMI
↓	↓	↓
non-specific mortality		death following AMI

Appendix 2 Adjustment factors for intermediate diseases

This appendix derives the formulas in sections 3.3.1.1 and 3.3.1.2 for the simple case of a single intermediate disease. The complete details how to deal with more complex relations between risk factors and diseases are given in a background report.[16]

First, some notation is introduced. We refer to Figure 8 in section 3.3.1.1 for illustration.
Let A denote with disease A, and $\bar{A}$ denote without disease A. Similarly, let B denote with disease B and $\bar{B}$ denote without disease B.
Let p_A, p_B denote the prevalence rate for disease A and B respectively. Denote $p_B(A)$ for the prevalence rate of B in those with disease A. Let $p_{\bar{A}}$, $p_{\bar{B}}$ denote the prevalence rates of $\bar{A}$ and $\bar{B}$, these equal (1- p_A) and (1- p_B).
Let Inc_C denote the incidence rate of disease C.

Define by $PRR_B(A)$ the disease A prevalence rate ratio of B, that is:
$PRR_B(A) = p_B(A)/p_B(\bar{A})$, or rewritten: $p_B(A)= PRR_B(A)*p_B(\bar{A})$,
$PRR_B(\bar{A}) = 1$ by definition.
The mean value of the disease A prevalence rate ratio of B all disease states of A can be written: $E(PRR_B)= p_A* PRR_B(A)+(1- p_A)*1$
It follows that $p_B =p_A*p_B(A)+(1- p_A)*p_B(\bar{A}) = p_A*PRR_B(A)*p_B(\bar{A})+(1-p_A)*p_B(\bar{A}) = E(PRR_B)* p_B(\bar{A})$

Define by $IRR_{C,unadj}(A)$ the disease A incidence rate ratio of C, (unadjusted)
Define by $IRR_C(A)$ the disease A incidence rate ratio of C, (adjusted for disease B)
Define by $IRR_C(B)$ the disease B incidence rate ratio of C, that is:
$IRR_C(B) = Inc_C(B)/Inc_C(\bar{B})$.
Assume this rate ratio is independent of the disease state for disease A, given B, so that
$IRR_C(B) = (Inc_C \text{ in } B|A)/(Inc_C \text{ in } \bar{B}|A) = (Inc_C \text{ in } B|\bar{A})/(Inc_C \text{ in } \bar{B}|\bar{A})$
This can be rewritten to
$(Inc_C \text{ in } B|\bar{A})= (Inc_C \text{ in } \bar{B}|\bar{A})* IRR_C(B)$, and $(Inc_C \text{ in } B|A)= (Inc_C \text{ in } \bar{B}|A)* IRR_C(B)$.

Define by $IRR_C(B)$ the disease B incidence rate ratio of C, that is:
$IRR_C(B) = (Inc_C \text{ in } B)/(Inc_C \text{ in } \bar{B})$.
Assume this rate ratio is independent of the disease state for disease A, given B, so that
$IRR_C(B) = (Inc_C \text{ in } B|A)/(Inc_C \text{ in } \bar{B}|A) = (Inc_C \text{ in } B|\bar{A})/(Inc_C \text{ in } \bar{B}|\bar{A})$
This can be rewritten to
$(Inc_C \text{ in } B|\bar{A})= (Inc_C \text{ in } \bar{B}|\bar{A})* IRR_C(B)$, and $(Inc_C \text{ in } B|A)= (Inc_C \text{ in } \bar{B}|A)* IRR_C(B)$.

Now, write down the unadjusted incidence rate ratio of A:
$IRR_{C,unadj}(A)= Inc_C(A)/Inc_C(\bar{A})$
$= \{(Inc_C \text{ in } B|A)* p_B(A)+ (Inc_C \text{ in } \bar{B}|A)*p_{\bar{B}}(A)\}/ \{(Inc_C \text{ in } B|\bar{A})* p_B(\bar{A})+ (Inc_C \text{ in } \bar{B}|\bar{A})*p_{\bar{B}}(\bar{A})\}$
$= \{ (Inc_C \text{ in } \bar{B}|A)*IRR_C(B) * p_B(A)+ (Inc_C \text{ in } \bar{B}|A) *p_{\bar{B}}(A)\}/ \{(Inc_C \text{ in } \bar{B}|\bar{A})* IRR_C(B) * p_B(\bar{A})+ (Inc_C \text{ in } \bar{B}|\bar{A})*p_{\bar{B}}(\bar{A})\}$
$= \{(Inc_C \text{ in } \bar{B}|A)/ (Inc_C \text{ in } \bar{B}|\bar{A})\}*\{ IRR_C(B) * p_B(A)+ p_{\bar{B}}(A)\}/\{ IRR_C(B) *p_B(\bar{A})+ p_{\bar{B}}(\bar{A})\}$
$= IRR_C(A)* \{ IRR_C(B) * PRR_B(A)*p_B(\bar{A}) + p_{\bar{B}}(A)\}/\{ IRR_C(B) * p_B(\bar{A})+ p_{\bar{B}}(\bar{A})\}$,
that is, the unadjusted rate equals the adjusted rate times a factor.

By assumption 9, $PRR_B(A) \approx IRR_B(A) \approx RR_B(A)$
Since $p_{\bar{B}}(\bar{A})= 1- p_B(\bar{A})$, $p_{\bar{B}}(A)$ equals 1- $PRR_B(A)*p_B(\bar{A})$, and $p_B(\bar{A})= p_B/E(PRR_B)$, the factor may be rewritten as follows:

$\{ IRR_C(B) * PRR_B(A)*p_B(\bar{A}) + p_{\bar{B}}(A)\}/\{ IRR_C(B) * p_B(\bar{A})+ p_{\bar{B}}(\bar{A})\}$
$=\{ RR_C(B) * PRR_B(A)*p_B + E(PRR_B) - p_B * PRR_B(A)\}/\{ RR_C(B) * p_B+ E(PRR_B) - p_B\}$
$=\{ E(PRR_B) +(RR_C(B)-1) * RR_B(A)*p_B \}/\{ E(PRR_B)+(RR_C(B)-1) * p_B\}$
$=\{p_A*PRR_B(A)+(1-p_A)+(RR_C(B)-1)*RR_B(A)*p_B\}/\{p_A*PRR_B(A)+(1-p_A)+(RR_C(B)-1)*p_B\}$
$= \{1+(RR_B(A)-1)*p_A +(RR_C(B)-1)* RR_B(A)*p_B \}/\{1+(RR_B(A)-1)*p_A+(RR_C(B)-1)* p_B\}$

Similarly, for the risk factor r with risk factor classes i (and class 0 the class with a risk ratio of 1),
define by $IRR_{C,unadj}(i)$ the class i incidence rate ratio of B, (unadjusted)
and by $IRR_C(i)$ the class i of risk factor r incidence rate ratio of C, (adjusted for disease B)
Define $PRR_B(i)= p_B(i)/ p_B(0)$
and write down
$IRR_{C,unadj}(i)= Inc_C(i)/Inc_C(0)$
This can be rewritten to
$\{ Inc_C(i)/Inc_C(0)\}* \{1+(RR_C(B)-1) (RR_B(i) 1)*p_B\} / \{E(RR_B)+(RR_C(B)-1)* p_B\}$

Appendix 3 Derivation of formula in section 3.3.1.4

This appendix derives the formula on page 29 for the case of two diseases. At first we describe the effect of competing death risks. As a result the unadjusted excess mortality rates are described as a function of the adjusted ones and the disease clustering. Then we introduce a risk factor. As a result the unadjusted excess mortality rates are described as a function of the adjusted ones and the other causes mortality rates. Finally we combine both models to describe the effect of dependent competing mortality risks with dependency through joint risk factors.

Notation:

A, E	disease indicator variables; $\bar{A}$: not having disease A
$p_E(A\|t)$	disease E prevalence rate conditional on having disease A
$m(e,a\|t)$	all cause mortality rates conditional on status a and e for disease A and E respectively
$em_A(t)$	unadjusted disease-related excess mortality rates
$em_{A0}(t)$	disease-related excess mortality rates adjusted for competing mortality risks
$em_{E0}(t)$	the disease E excess mortality rate adjusted for disease A,
$p_{A,E}(t)$	the prevalence of diseases A and E simultaneously

The adjusted disease-related excess mortality rates may be written as:

$$em_{A0}(t) = m(A,E|t) - m(\bar{A},E|t) = m(A,\bar{E}|t) - m(\bar{A},\bar{E}|t)$$
$$em_{E0}(t) = m(A,E|t) - m(A,\bar{E}|t) = m(\bar{A},E|t) - m(\bar{A},\bar{E}|t)$$

The unadjusted disease A excess mortality is defined as: $em_A(t) = m(A|t) - m(\bar{A}|t)$ and this maybe rewritten to:

$$em_A(t) = m(A,E|t)\, p_E(A|t) + m(A,\bar{E}|t)\, p_{\bar{E}}(A|t) - m(\bar{A},E|t)\, p_E(\bar{A}|t) - m(\bar{A},\bar{E}|t) p_{\bar{E}}(\bar{A}|t)$$

$$= em_{A0}(t) + em_{E0}(t)\, \{\, p_E(A|t) - p_E(\bar{A}|t)\, \}$$

adjusted disease A excess mortality + *excess mortality through disease E*

with:

$$p_E(A|t) - p_E(\bar{A}|t) = \{\, p_{A,E}(t) - p_A(t)\, p_E(t)\, \} / \{\, p_A(t)\, (1-p_A(t))\, \}$$

here $p_{A,E}(t) - p_A(t)\, p_E(t)$ may be interpreted as the disease A and E excess co-morbidity rate, that is, the additional co-morbidity in comparison to the comorbidity that will always exist even if the two diseases are independent.

Summing up:

$$em_A(t) = m(A|t) - m(\bar{A}|t) = em_{A0}(t) + em_{E0}(t)\, \{\, p_{A,E}(t) - p_A(t)\, p_E(t)\, \} / \{\, p_A(t)\, (1-p_A(t))\, \}$$

If this same line of reasoning is followed for the case of more than two diseases, the formula in section 3.3.1.4 results.