TECHNISCHE UNIVERSITÄT WIEN

VIENNA UNIVERSITY OF TECHNOLOGY

DIPLOMARBEIT

A System Dynamics Model for the Diabetes Prevalence in Austria

ausgeführt am

Institut für Analysis und Scientific Computing

der Technischen Universität Wien

unter der Anleitung von

Ao. Univ. Prof. Dr. Felix Breitenecker

durch

Dipl. Ing. Peter Kristöfel

Heiligenstädterstraße 11-25/14/12

1190 Wien

Datum

Unterschrift

Zusammenfassung

Typ-2 Diabetes Mellitus ist auf dem Vormarsch in unserer alternden Wohlstandsgesellschaft. Als chronische Krankheit mit schwerwiegenden Folgen stellt sie eine große Herausforderung an das Gesundheitswesen dar. Die Frage ist, wie am besten mit dieser ernsten Bedrohung der Volksgesundheit umgegangen werden kann. In dieser Diplomarbeit wird ein System Dynamics Modell zur Simulation der Diabetes Prävalenz in Österreich entwickelt.
Die vielen verschiedenen Eingangsdaten sind von unterschiedlicher Qualität, sie ändern sich diskontinuierlich mit der Zeit. Die statistische Datenerfassung erfolgte selektiv, nur für einige Merkmale zu unterschiedlichen Zeitpunkten, jedoch sind nicht alle notwendigen Parameter verfügbar. Zum einen ist durch diese Einschränkungen ein rein statistisches Modell unmöglich, zum anderen wirken sich diese Beschränkungen auf die Struktur des dynamischen Modells aus. Die Modellierung des Krankheitsverlaufs in der Bevölkerung führt also auf ein System von gekoppelten, nichtlinearen, algebraischen Integro-Differentialgleichungen mit diskreten Zustandsänderungen im Zeitverlauf.
Dies legt eine Behandlung mit System Dynamics nahe, moderne System Dynamics Programme ermöglichen die Berechnung der zeitlichen Entwicklung eines solchen Systems. Ein von J. Homer et al. [1] für die USA entwickeltes Modell wird an die Struktur des österreichischen Gesundheitssystems und an die verfügbaren Daten angepaßt und erweitert. Insbesonders wird eine Unterscheidung nach dem Geschlecht eingeführt, denn es besteht ein Geschlechter spezifisches Risiko an Typ-2 Diabetes zu erkranken.
Die verfügbaren Eingangsdaten werden so in das Modell implementiert, dass die korrekte historische Entwicklung der Verbreitung von Diabetes in Österreich reproduziert wird. Die Stabilität des Systems wird mit statistischen und Monte-Carlo Methoden untersucht. Danach werden einige Experimente, angelehnt an derzeit laufende Studien, jedoch mit einer größeren fiktiven Population, mit dem Modell durchgeführt. Parameter, mit denen der Erfolg von unterschiedlichen Maßnahmen evaluiert werden kann, werden identifiziert.
Ein zukünftiger Erweiterungsschritt könnte die Anwendung des Modells auf Personen mit unterschiedlichem sozialen Hintergrund sein oder die Verbindung mit einem Modell zur Verbreitung von Adipositas, da eine starke Korrelation besteht.

Abstract

Type-2 diabetes mellitus is on the advance in aging affluent society. As a chronic disease with severe consequences it poses a major health care challenge. The question is how to best manage this serious threat to public health. In this thesis we develop a System Dynamics model for the type-2 diabetes prevalence in Austria.
There are many different input variables and they change discontinuously with time. The statistical surveys to obtain these data were performed selectively, just for some characteristics at different times, and not all parameters are available. On the one hand this makes a model relying only on statistics impossible, on the other hand these restrictions influence the structure of the dynamic model. So the modelling of the course of the disease in the population leads to a system of coupled, nonlinear, algebraic integro-differential equations with discrete changes in state with time.
Therefore System Dynamics is the method of choice. Modern System Dynamics programs allow the calculation of the time development of such a system. A model developed by J. Homer et al. [1] for the USA is adopted to the health care system in Austria and the available data. It is enhanced further, especially a distinction by sex is introduced, since there is a gender-specific risk to develop type-2 diabetes.
The available input data is implemented in the model to reproduce the correct historic prevalence of diabetes in Austria. The stability of the system is examined with statistical and Monte-Carlo methods. Then some experiments with the model, analogous to ongoing studies but with a larger population, are made. Parameters with which the success of different measures can be evaluated are identified.
Future work may include the extension of the model to people from a different social background or the connection with a model for the adiposity prevalence, since there is a strong correlation.

For my grandmother, Anna Lindmayr,
who suffers since 1992 from type-2 diabetes.

Contents

List of Figures

Chapter 1

Introduction

Diabetes mellitus and its complications are one of the most challenging topics in public health care. It is a chronic, progressive and fast-spreading disease. The World Health Organization (WHO) [2] estimates that over 180 million people worldwide have diabetes. In 2005 an estimated 1.1 million people died from diabetes. Taking into account deaths in which diabetes was a contributory condition this number increases to about 2.9 million people. Almost 80% of diabetes deaths occur in low and middle-income countries. Almost half of diabetes deaths occur in people under the age of 70 years and 55% of diabetes deaths happen to women. The WHO expects a 50% increase of the diabetes deaths for the next 10 years.

In the European Union 25 million people are known to have diabetes [3]. It is assumed that 50% of the people with diabetes are not aware of their condition, so that over 50 million people may be affected. And 60 million people are at risk of developing pre-diabetes. Diabetes is therefore a major chronic disease responsible for 5 to 10 percent of the total health care costs. In the EU estimated 30 billion Euro are spent each year on the treatment of diabetes and its consequent illnesses. The WHO expects a rise in diabetes prevalence of 37 percent from the year 2000 to the year 2025. So the question is how to manage this serious threat to public health in the best way.

One powerful tool to help finding the best strategy is System Dynamics. With such models it is possible to analyze the time dependent behavior of all kinds of systems. Successful applications have been made in economy [4], engineering, medicine ([5], [1]) and politics.

The structure of this thesis is as follows: in chapter one we will give the basic

informations about diabetes mellitus and its consequent diseases. We also introduce the basics of System Dynamics and describe its applicability to model health care problems. In chapter two the System Dynamics model used to describe the spread of type-2 diabetes will be discussed. Chapter three lists the input parameters as well as their sources. The model is analyzed and validated in chapter four. Various experiments with different health care policies are tested and interpretations are given. A summary of the results and an outlook to future work is given in chapter five.

1.1 Diabetes Mellitus

Diabetes Mellitus is a metabolic disease which is known since the ancient Greek. The name of the disease is deduced from the Greek $\delta\iota\alpha\beta\acute{\eta}\tau\eta\zeta$, which means "passing through" or "siphon", and the Latin *mellitus* meaning honey [6]. These terms refer to two major symptoms of diabetes mellitus: excessive urine production with a sweet taste due to an excess of one kind of sugar in the urine. Ancient Indians as well as the Korean, Chinese and Japanese also call it the "sweet urine disease". The reason for these symptoms is a disturbance of the glucose metabolism which leads to a high blood sugar level (hyperglycemia). This is due to either the lack or the ineffectiveness of the anabolic ("building up") hormone insulin.

Usually most of the carbohydrate taken in with food is cracked down to monosaccharide glucose (a form of sugar) by the digestive system. Carbohydrates which are not converted include fruit sugar (fructose) and cellulose. While fructose is passed on into the blood stream and used as a cellular fuel, cellulose can't be digested by most animals, including the human. The monosaccharide glucose is transported through the intestinal wall into the blood and distributed throughout the whole body. In response to an increase of blood sugar level the beta-cells in the islets of Langerhans of the pancreas start releasing insulin. Insulin helps most of the body cells, like muscle cells and adipose tissue, to take up glucose from the blood stream into the cell. There it is used to fuel the cell or to synthesize other needed proteins. Insulin is also the trigger to convert glucose to glycogen, which in turn is stored in muscle or liver cells. Also the production of glucose in the liver is inhibited by insulin. The third effect of insulin is that it is necessary for the built up and storing of body-fat. If the production of insulin is disturbed or the cells can't process the insulin right,

the blood sugar is not transported into the cells and the liver produces even more glucose. Hyperglycemia and the consequent symptoms describing diabetes mellitus occur.
The WHO distinguishes between five types of diabetes mellitus [7]. After the International Statistical Classification of Diseases and Related Health Problems standard in its current version (ICD-10) there are:

- **E10:** Insulin-dependent diabetes mellitus (type I diabetes)
- **E11:** Non-insulin-dependent diabetes mellitus (type II diabetes)
- **E12:** Malnutrition-related diabetes mellitus
- **E13:** Other specified diabetes mellitus
- **E14:** Unspecified diabetes mellitus

Type I diabetes, also called childhood diabetes or juvenile diabetes, encompasses all diabetes cases with absolute insulin deficiency. Here the beta-cells are destroyed, most of the time through an autoimmune reaction, leading to a relative insulin deficiency. It is responsible for around 10% of all diabetes cases, although this number may vary geographically between 5% in Asian and 15% in Scandinavian countries [8]. It usually starts in the early childhood and there is a strong genetical predisposition.
About $90 \pm 5\%$ of all diabetes cases are of type-2. Other names are aquired diabetes, maturity-onset diabetes or obesity related diabetes. In this case the pancreas is still capable of producing insulin, but the cell-membranes have a reduced sensitivity or even complete resistance to it. Since this will be the subject of our study we will describe it below in more detail.
Other types include diabetes due to genetic defects, diseases of the pancreas, chemicals or drugs and gestation diabetes. Their total prevalence is negligible compared to type-1 and type-2 diabetes.

Type-2 Diabetes Mellitus

The following facts about type-2 diabetes mellitus are a collection of data mainly from the references [9], [10] and [11].

In an early stage of type-2 diabetes the cell membranes have a reduced insulin sensitivity, which is compensated by an increased insulin level. As the disease progresses the pancreas can't keep up to produce enough insulin, leading to hyperglycemia and the type-2 diabetes becomes manifest. But it may still be years before it is diagnosed.

Causes

There is a great variety of theories of the exact causes and mechanisms leading to insulin resistance. One of the main risk factors is obesity, especially central obesity, since abdominal fat is hormonally very active. With increasing age there is also a much higher risk of developing type-2 diabetes. In recent years also younger people are affected. In low- and middle-income countries most of the diabetes cases occur in the middle-aged group of 45-64 years and not in the elderly group (65+). This supports the theory that type-2 diabetes mellitus is an acquired chronic disease. There is also a strong genetical influence. A hereditary insulin resistance makes sense for people living in an environment where they are subject to food shortage. A high insulin level ensures that all of the carbohydrates taken in are used metabolically. Already in childhood people with this genetical predisposition tend to become overweighted. Especially in threshold countries, where there is bumper food supply, this leads to an explosive increase in type-2 diabetes. Other risk factors are the lack of physical activity, smoking, alcohol consumption and an unhealthy diet. These are usually summarized under the term 'lifestyle factors'. Low weight at birth, gestation diabetes and socioeconomic factors are also rick factors in developing type-2 diabetes.

Aside from the aforementioned causes there is also the metabolic syndrome. It is regarded as a precursor of type-2 diabetes, but there are several definitions around. All include a large waist circumference, hypertension (high blood pressure), high triglyceride levels or low HDL ("good") cholesterol and high blood glucose levels. However, the term metabolic syndrome has been criticized since it has no unique definition and an unreflected use would pathologise great part of the population. We will therefore use the term prediabetes instead, which describes the state before type-2 diabetes. Here the blood sugar level and the insulin production are already increased, but not as high as fully developed diabetes. In this stage the symptoms can still be reverted.

Diagnosis

The symptoms of type-2 diabetes are rather unspecific: tiredness, weakness, permanent hunger, blurred vision, weight gain and depressions. It may be years till it is diagnosed since the symptoms are mild or even nonexistent and they fit to any other disease as well. Still the undiagnosed disease may lead to severe but characteristic complications (c. f. below). Most diagnosis come from health screenings or after the recent onset of one or more symptoms or consequent diseases [12].

The medical criterion is an, at least twice measured, elevated blood plasma glucose level. There are two ways of testing: Firstly there is the fasting plasma glucose test (FPGT), which indicates diabetes if the plasma glucose level is at or higher than 126mg/dl (6.1mmol/l). Secondly there is the oral glucose tolerance test (OGTT), which indicates diabetes when the plasma glucose level is at or higher than 200mg/dl (11.1mmol/l) two hours after the consumption of 75g glucose. Obtaining just one positive test result together with one or more of the characteristic symptoms is also considered as a positive diabetes test.

Both tests can also be utilized to diagnose a disturbed glucose tolerance (prediabetes). The respective limit values are 110-125mg/dl for the FPGT and 140-199mg/dl for the OGTT.

Complications

There are three kinds of acute complication associated with diabetes. The first is hypoglycemia, or abnormally low blood glucose. This is the consequence of either an overdose of blood sugar reducing medicaments or eating to little while under treatment of such. The symptoms vary individually but are similar to a severe hangover. The reason is that alcohol metabolization in the liver prohibits the production of glucose there. The second acute complication is diabetic ketoacidosis, when no insulin is present in the blood. The consequence is that, although the blood sugar level is high, the liver starts the production of ketone bodies, another form of fuel for the cells. This is part of the fat metabolic processes, but these ketone bodies are not needed. Since they are positively charged they lead to a low blood pH value, which is an acidic poisoning. The symptoms are dehydration and a high level of stress hormones. The third acute complication, the hyperosmolar nonketotic state, has similar symptoms but a different cause. The reason is the high blood sugar

level. Water is drawn osmotically out of the cells into the blood. Additionally the kidneys try to get rid of the excess blood sugar with the urine. This is done with much water and the kidneys can't resorb enough of it. So the osmotic pressure of the blood increases further and additional water is drawn out of the cells. This leads to dehydration.
The most common accompanying disease of diabetes is hypertension (about 75% of all cases). Since chronic hyperglycemia damages the blood vessels consequent diseases include the following: Retinopathy due to the growth of friable and poor-quality new blood vessels, which can lead to severe sight loss or even blindness (2% after 15 years). Neuropathy due to the damage of peripheral neurons, which result in the alteration or loss of feeling, especially in the feet. Coronary artery disease, which leads to angina ("heart ache"), a myocardial infarct ("heart attack") or a stroke. Over 50% of the people with diabetes die due to it. Peripheral vascular disease, which increases, togehter with the neuropathy and circulatory disorders, the chance of foot ulcers and eventual limb amputation (these symptoms are often called a diabetic foot). Muscle wasting may occur (myonecrosis), as well as gum bleeding and parodontitis. And finally the kidneys may be damaged (nephropathy), requiering the need of dialyses. Diabetes is among the leading causes for kidney failure. Overall the risk of dying is at least twice as high compared to people without diabetes.

Therapy

Diabetes mellitus is a major chronic disease of our aging, affluent society. Currently there is no cure. Therefore the medical treatment must focus on the management and the prevention of possible short- and long-term complications. In many countries a diabetes plan has been decided and disease management programs are implemented [13]. Important parts on the patient side are: education, a diabetic diet, more physical activity, self glucose monitoring, regular controls and, if indicated, various oral diabetic medicaments or insulin treatment. For the care staff high quality training and education has to be assured. A sufficient number of care institutions and enough money has to be available. And the data has to be collected and evaluated regularly in a suitable fashion. Always all accompanying diseases, foremost hypertension, have to be watched and treated appropriately. Screening and early detection help reducing the burden of diabetes and it's consequences.

Prevention

In a scientific context three levels of prevention are defined:

- Primary prevention which avoids the onset of a diabetes by a reduction of risk factors. This includes measures for high-risk patients, prophylactic measures for the whole population and programs for specific target groups (social surrounding, women, children, ...) or specific settings (at work, school, university, ...).
- Secondary prevention aims to avoid retinopathy, nephropathy, neuropathy, etc. after the diabetes onset.
- Tertiary prevention aims for preventing the loss of functions, such as blindness, amputation, the necessity of dialysis or kidney transplantation, etc..

Secondary and tertiary prevention can try to reduce risk factors or can target the disease itself. An overview of preventive measures for type-2 diabetes at all levels in Austria and the EU is given in the references [3] and [13].

1.2 System Dynamics

Since the development of computers simulation has become an indispensable aid in science and technology. Under simulation we understand the aimed experimenting with models which mimic a part of the real world or a simplified equivalent. It is used since the experiments would be too expensive and time consuming or could not be done at all (e. g. in macroeconomics). One very successful simulation method, which we will apply to the study of type-2 diabetes mellitus, is System Dynamics (SD).

It has been invented by Jay W. Forrester in the late 1950s at the Massachusetts Institute of Technology Sloan School of Management. Originally called 'Industrial Dynamics' it was designed to investigate dynamical models in the industry. Soon afterward the method was applied in many other fields, ranging from economy [4] and business [15] over physical systems to politics and even world models [16]. It is a method for the simulation of complex dynamical systems, to study long term effects and compare different policies. The aim is to support decision making processes when other tools, like operations research, are not applicable. Another advantage

is that SD is quite intuitively. So experts from different fields of study with highly specialized knowledge can work together on a project. The simple main parts out of which System Dynamics models are built are stocks, flows and feedback loops. These are introduced now in the following paragraphs.

Systems

The starting point of SD is a system, which is defined as a collection of elements together with the interconnections between these elements. Each element has at least one connection to another element and all elements are connected. Through this, a system has a well defined border which separates it from the environment. This defines automatically two levels: the system as a whole (macro-level) and the elements as parts of the system (micro-level). Note that these elements may be systems itself, which are then called sub-systems. The interaction between the elements makes the whole more than just the sum of its parts. Systems have different qualities. All systems can be assessed under the following aspects (see e. g. reference [14]):

- Dynamical systems exhibit a time dependent behavior as opposed to static systems.
- There are discrete systems and their opposite continuous ones.
- Open systems interact with their environment (or some elements may change) in contrast to closed systems.
- There are deterministic and stochastic systems, weather the starting quantities and calculations are sharp or random in nature.
- Nonlinear systems may exhibit chaos. This also gives rise to the concept of the stability in this context.
- The degree of complexity depends on the number of elements and connections as well as the strength of the influences of each component.

There are further properties which are present in some systems, like autonomy (independence from the environment), self-reference (no input needed), the abilities to learn, self-regulation and -control, adopt, think, ... Many of these terms have no exact definition since not all systems are comparable.

As one short example lets take a look at a car: the engine may be seen as an element of the system car. The engine consists of many sub-systems, like spark plugs, valves, etc. . Since the engine can perform work it is more than the sum of its parts, as well as the car, which can hopefully drive. The car is an open system since it needs fuel. It is dynamic since it can move onward and the internal states are altered periodically in time, which in turn is a kind of stable behavior. The mechanics is rather simple, in contrast to modern car electronics, which may be self-monitoring. The fuel-injection system together with all relevant components are self-controlling. But surely a car is not thinking, at least without high level navigation computers. A comparison to the human brain or the stock exchange, which are other highly complex dynamic system, with the same classification categories is impossible. System dynamics is an analysis method suited for time dependent complex systems which are usually nonlinear.

Cause and Effect Diagrams

The main characteristic of System Dynamics is the occurrence of feedback loops. These result from the causal structure of the system. The connections of the system are analyzed whether one element has a positive or a negative influence on another element in the sense of cause and effect. The respective sign is attributed to the connection. All elements and connections may be drawn into a causal-loop-diagram. It can now happen that one element has, via a detour over other elements, an effect on itself. What we get is a feedback loop. Typically SD models have many feedback loops leading to nonlinear and counterintuitive behavior. One example for this are the occurence of Nash-equilibria in game theory when the participants of a game start to think of what the others would do [17]. From a systems point of view three main types of feedback can be distinguished: positive feedback leading to growth and possibly escalation, negative feedback leading to stabilization and indifferent feedback if a positive and a negative loop are present for the same element. To determine if a loop is positive or negative one has just to multiply the signs of all connections which are part of the loop. For mechanical systems and electrical circuits the drawing of causal-loop-diagrams can be formalized to the bond-graph method. A nice example for a positive feedback loop are self-fulfilling prophecies [21]. Another is population growth with unlimited resources, as seen in figure 1.1.
A simple negative feedback is the demand-price relationship, as pictured in figure

1.2: the more demand the higher the price. And a higher price leads to less demand, which in turn lets drop the price. In total the loop has a negative sign and we have a stabilizing effect. But if there are many feedback loops the general behavior may depend very sensitively on the exact numerical values.

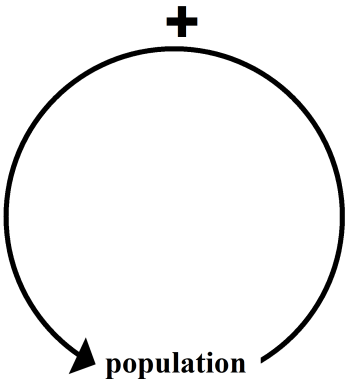

Figure 1.1: Positive feedback loop: an increase in population leads, via more births, to even more population.

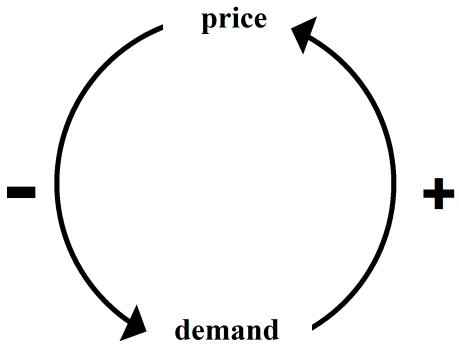

Figure 1.2: Negative feedback loop: higher demand leads to higher prices, but higher prices to a lower demand.

In the development of a SD model it is necessary to think of all possible influencing factors. In a first step all of them have to be included, since the feedback loops may enhance the influence of a small factor significantly. Therefor the analysis of the causal structure and the finding of feedback loops is of uttermost importance. The logical configuration sets what can be done with the model and which effects can be produced.

Stocks and Flows

The active elements of a System Dynamics model are the stocks (levels). They describe the filling level of reservoirs, where the quantity one is interested in can be accumulated, stored in or expended from. These stocks are connected by flows, which are the rates of change. While the stocks usually just integrate the net in- and outflow rate, the flows may be difficult to calculate and may depend on many

variables[1]. When the time starts to run we see rising and falling levels and springing and ebbing flows with each time step.
Some examples for flows and stocks are: the incoming and outgoing commodities in a warehouse, the in- and outflowing liquids in tanks, the electrons in a capacitor, ... For population levels the flows may be births and deaths (c.f. figure 1.3), immigration end emmigartion, arrival and departure or infections and healing. A stock and flow diagram can be drawn after the causal-loop diagram has been found. It is the first step to a quantification of a model.

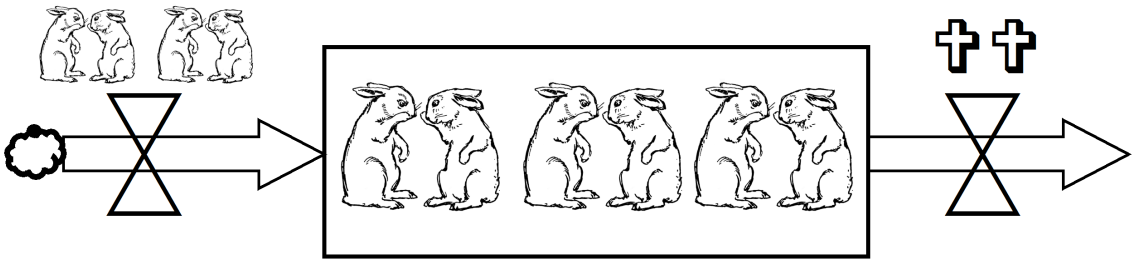

Figure 1.3: The rabbit population level: small baby bunnies are born into and old die out of the population.

Equations

After the causal-loop diagram and the stock and flow diagram have been found the connections and flows need to be quantified. This is done by finding the corresponding mathematical equations. They are usually integro-differential algebraic equations with suitable initial conditionswhich are not analytically solveable. Here the whole mathematical framework of solving such equations numerically can be applied. Fortunately modern simulation programs have graphical interfaces and the simulation engine takes care of the numerics.
A feature often encountered is the occurrence of time-delays. This means that the rates with which the flows transport stuff from one level to another respond delayed to changes. An example is an electronic circuit consisting of a voltage source and a capacitor: when the voltage source is turned off, the voltage will start to drop but not immediately jump to 0. Since long time delays can be modeled, System

[1]We use the following terminology: a parameter does not depend on time while a variable may be time-dependent.

Dynamics is useful for studying long-time effects. This is the reason why it is a major tool of strategic controlling [15].
After having found the causal structure and quantified the functional relationships one can test the model if it produces useful and realistic results. If this is not the case one has to start the model building cycle again and to find points where improvements can be made.

1.3 Modelling Health Care Problems

Besides profane statistics different alternative modelling approaches for health care problems are available: time-series analysis and Markov-chains [18], agent based models [19], cellular automata [20], genetic algorithms, stochastic processes, Mont-Carlo methods, System Dynamics and many more. Depending on the specific questions and settings one or more of them may be appropriate. In the following we will give some arguments why System Dynamics is the number one choice for modelling the type-2 diabetes prevalence.
Chronic diseases are widespread in our aging affluent society. Approximately one third of the total population suffers from chronic diseases and with increasing age the percentage is rising steeply. Already more than one half of the people above 60 suffer from at least one chronic disease. Therefore these afflictions are responsible for a great part of the health care costs, outstripping the costs of accidents and acute diseases combined. According to all estimates chronic diseases are going to increase further.
Besides the socioeconomic importance, chronic diseases sport several features which suggest a System Dynamics treatment:
Firstly all health care officials, including doctors, politicians, patient associations and other medical staff, recognize the threat and agree that measures on an ecological, system-wide level have to be taken to reduce chronic diseases and their consequences. But most programs sport conventional analytical methods by which each aspect of a complicated disease control strategy is addressed and evaluated separately. The advantage of System Dynamics here is that one gets a global picture where all influencing factors are incorporated and act together.
Secondly chronic diseases involve long time scales. There are long delays between causes and health consequences making short term analysis methods unsuitable.

The third argument why System Dynamics is suitable is: for every one of the three prevention levels many different policies are available. Primary prevention includes behavioral and socioeconomic measures like improving lifestyle, working and living conditions, information, education and many more. Secondary prevention focuses on precaution and early detection. And finally elements of the tertiary prevention are accessibility to the medical treatment, improvement of compliance and empowerment. All these measures together with quality control are elements of a process-based disease-management approach. System Dynamics now gives the opportunity to test different approaches and policies simultaneously and observe the respective outcome.

Also the intuitive character of SD makes it easy for experts from different fields to work together interdisciplinary.

Finally diabetes mellitus is the prime example of a chronic disease. Together with it's many different causes, prevention and therapy methods as well as malign consequences diabetes forms a complex system. It is researched well enough so that the main risk factors are known and that much input data is available.

Chapter 2

The Model

In the following we will give a description of our model as well as the ideas behind structuring it as it was done. The basic structure was taken from the Diabetes System Model of Jack Homer et al. ([1], [22], [23]), which has been developed for the US-American Center of Disease Prevention and Control. The starting idea is to simulate the whole population of a given region as it moves through the different stages of type-2 diabetes. In this case a region may be a geographical area, a social surrounding or any other sufficiently large and homogeneous group. Since DM is not infectious it is irrelevant where the people belonging to such a group move around geographically. SD is after all a top down approach, where the laws of evolution for the whole population are laid down and where it does not matter with whom individuals interact. This makes the modelling life a bit easier.

There are three main factors influencing type-2 diabetes: firstly adiposity, through the caloric balance per day, which is affecting the average body-mass-index (BMI). Secondly the fraction of elderly people, who have a significantly higher risk for diabetes. And thirdly the disease management, which includes testing frequencies, the access to health care institutions and benefits as well as the ability to self-management.

After introducing the main stock-and-flow diagram we will give a detailed analysis of the different in- and outflows as well as the transition rates. The adiposity feedback-cycle will be described separately, since only the average BMI enters the rest of the system. Here lies also one of the possible future enhancements of the system, since obesity is also a fast spreading disease with high costs and severe consequences in our society and much research is done on it [24]. The health care costs will be

examined separately, since these follow out of the model and have no feedback to it. An enhancement here could be to include the progression of the social security contribution and its influence on the provided health care measures.

2.1 Stocks and Flows

Figure 2.1 shows the population levels and flows of the model. The population levels are arranged in four groups: the first group consists only of one level, the people with a normal blood sugar level (normoglycemic population). The other three groups are the people with an elevated blood sugar level (prediabetes population), the people with diabetes and the people who have diabetes as well as an additional consequent disease or acute complication. These three groups all consist of two levels: the diagnosed and the undiagnosed ones. This distinction makes sense since there is a very large estimated number of unreported cases. It is absolutely necessary to incorporate these in this model: on the one hand screening and prevention methods could not be tested otherwise. On the other hand severe complications and even death may occur before type-2 diabetes mellitus is even diagnosed.

The four different groups have a connection to the three prevention levels: the first two groups are the targets of primary prevention, the prevention of the onset of a disease. This is especially true for the people with prediabetes since they are the risk group. Some people are easily identified as belonging to the risk group, like obese ones, but others may be difficult to recognize, like genetical predispositioned ones. The third group with fully developed diabetes is the target of secondary prevention, which aims to avoid deterioration. And finally the fourth group is the target of tertiary prevention, which aims to counteract the impeding loss of function. All levels are just the integrals of the net in- and outflow. Each of them has a finite starting value.

Let us now examine the flows more closely: all flows in the model are expressed as people per year. There is only one inflow of healthy adults into the fist level, while people may die out of every level. This inflow is given as a time series input by statistical predictions [28]. The resons why just the adult population is simulated are discussed in the next section where the actual rate will be calculated. The different death rates are affected by the fraction of elderly people, which is again given as a time series. For the diabetes population with complications the death

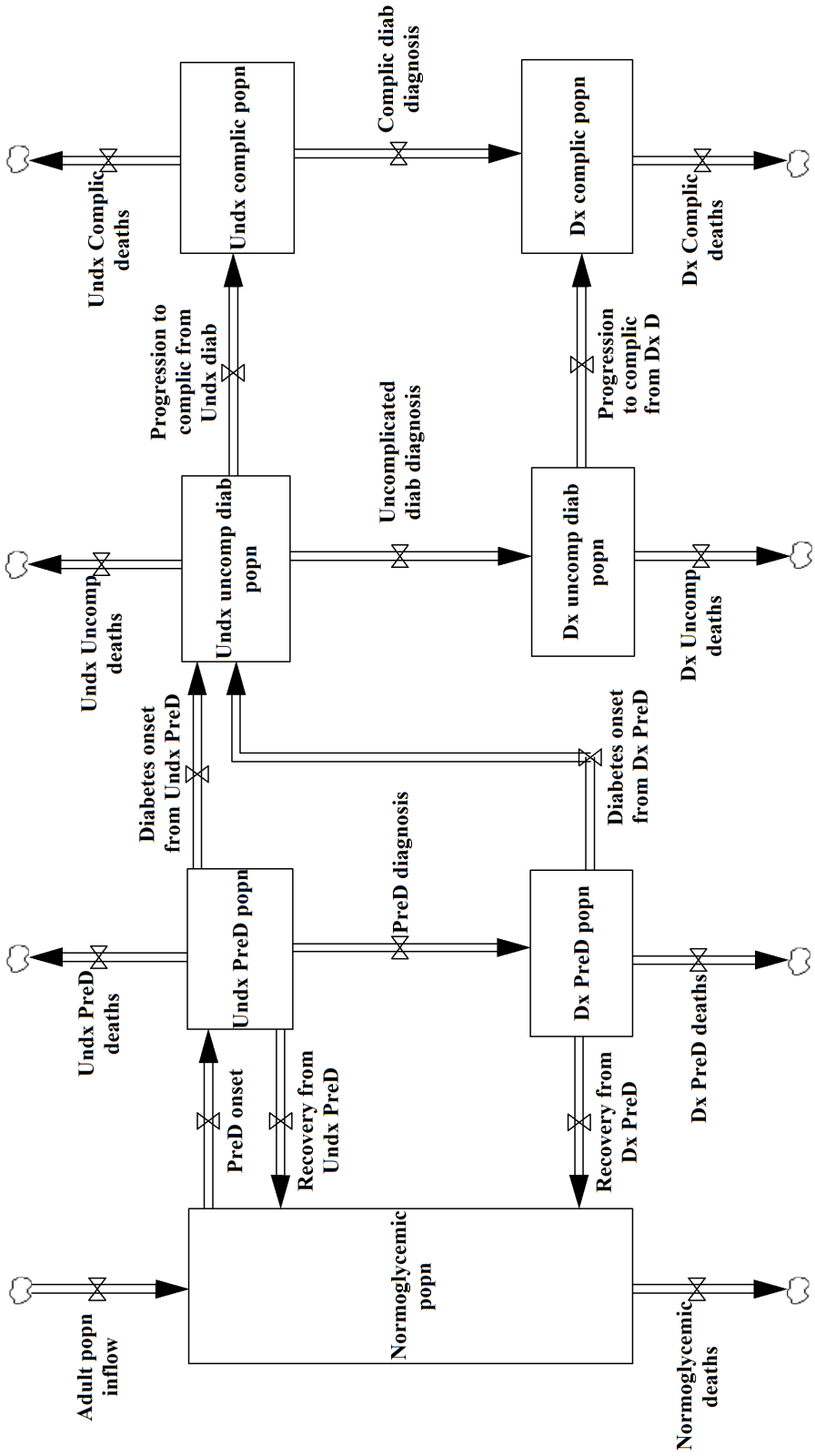

Figure 2.1: The population stocks and flows of the model.

rate is also influenced by these consequent diseases. The basic assumption is that the relative rates R of people with a risk factor E compared to people without it remains constant in the concerned group. The corresponding equation is

$$R = \frac{\mathbb{P}(d|E)}{\mathbb{P}(d|\neg E)} = \text{const.} \quad ,$$

where $\mathbb{P}(d|.)$ is the conditional probability of an attribute d to occur under the influence of the respective factor. This leads to the following transformations:

$$\begin{aligned} \mathbb{P}(d) &= \mathbb{P}(d|E) \cdot \mathbb{P}(E) + \mathbb{P}(d|\neg E) \cdot \mathbb{P}(\neg E) = \\ &= (R \cdot \mathbb{P}(E) + \mathbb{P}(\neg E)) \cdot \mathbb{P}(d|\neg E) = \\ &= ((R-1) \cdot \mathbb{P}(E) + 1) \cdot \mathbb{P}(d|\neg E) \end{aligned} \tag{2.1}$$

since $\mathbb{P}(E) + \mathbb{P}(\neg E) = 1$. With a suitable initial value for either $\mathbb{P}(d|\neg E)$ or $\mathbb{P}(d)$, the development of $\mathbb{P}(d)$ in dependence of $\mathbb{P}(E)$ can be calculated. Complications also increase the death rate for the respective levels. The death rate for the people with complications is also influenced by the control of the disease, the "disease management". With the initial values the dynamic death rates can then be calculated. The death rates are not constant ! When the population gets older the death rates increase. And also the death rate due to complications of DM can be influenced, which is the aim of the whole model.

The flows between the different stocks are characterized by the following assumptions: while people with prediabetes can still recover, there is no way to cure diabetes mellitus after its onset. Type-2 diabetes is a chronic disease and once complications occur the damage is dealt and until now no way of repairing it has been found. The onset of prediabetes and diabetes occur in principle unobserved, while complications can also arise even if under medical supervision. All transition rates from one level to another are affected by the elderly and the obese fractions of the respective populations. The progression rates in figure 2.1 (the horizontal ones) of the detected populations can be influenced by the clinical management, like prevention measures and compliance. The detection rates (the vertical ones) are more difficult to describe: they are first order exponentially delayed functions of the progression rates as well as the testing frequency and the sensitivity of the tests. Time dependent input data enter in several places of diabetes mellitus detection and control incorporating different possible health policies. The feedback loops and the details of the calculation of the flows are given in the following sections.

The color code for all graphs following is always the same: black are variables which are calculated in the model. Red characterizes constant input parameters. Pink denotes input parameters varying with time and output variables are orange. Finally gray variables in brackets are place holders and are defined at another point in the model.

2.2 Life and Death

In figure 2.2 we see the causal structure for the inflow and the deaths of the healthy population. There is the adult population inflow, which consists of all peoples reaching the age of 20 years. Although in recent years more type-2 diabetes has been detected in juveniles [11], we do not include them in the calculation, since their contribution with 0.25 out of 100000 cases (only 8 diagnoses in the years 1999-2001 in Austria) is negligible. So all people entering the model are healthy, from the viewpoint of diabetes at least, although they might be obese (c. f. section 2.5). The adult population is given as a time series, from which the growth rate is calculated straightforward. The adult population inflow is then the sum of the growth and the death rate for the adult population multiplied by last years Adult population. So this is the number of people turning 20 years of age.

The death rate for the adult population is the product of the initial death rate and an effect of the aging of the society. The effect of the population aging on the nondiabetes population death rate is calculated as in equation 2.1 from the elderly fraction of the adult population and a risk factor for the death rate for elderly compared with non-elderly adults. The death rate for the people without diabetes is calculated in the same way as for all adults with the only difference that the initial death rate is a little bit smaller. This death rate is the same for the normoglycemic population level and the prediabetes population levels since prediabetes is assumed to be nonlethal. The actual flow in deaths per year is then the product of the current level times the respective death rate. (All graphs in this section are best read backwards from the absolute number of deaths per year.)

The causal structure for the death rates for the population with diabetes can be seen in figure 2.3. The death flows out of the levels are calculated in the same way as for the non-diabetes population. The only difference is that the risk factor and the initial death rate are different.

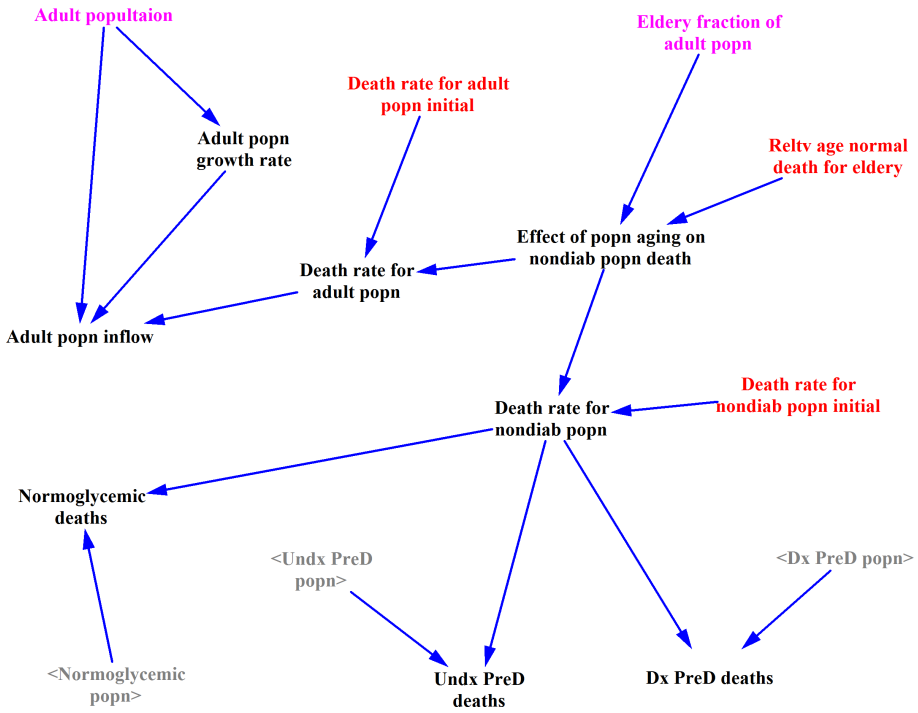

Figure 2.2: The causal structure for non-diabetes inflow and deaths.

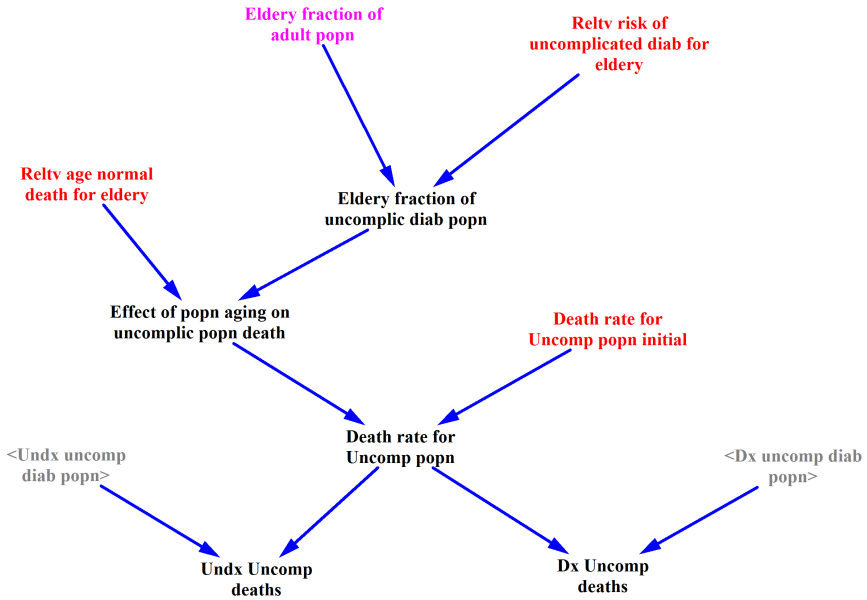

Figure 2.3: The causal structure for the uncomplicated diabetes deaths.

The situation for the population with complications is different, since the complications may lead to a premature death (especially through cardiovascular diseases, c. f. section 1.1). The causal structure, and therefore the equations, for the age-normal death rate for the complicated diabetes population is the same as in the previous two cases, as can be seen on the left hand side in figure 2.4. To this age-normal rate a death rate due to the complications is added. This rate due to complications is reduced if the diabetes is diagnosed and it is reduced even further if it is diagnosed and controlled:

> Complications death rate for Dx Complic = Complications death rate for Undx Complic*(Controlled fraction of Dx diabetes popn*Relative risk of complicatios death if controlled + (1-Controlled fraction of Dx diabetes popn)*Relative risk of complications death if Dx but uncontrolled)

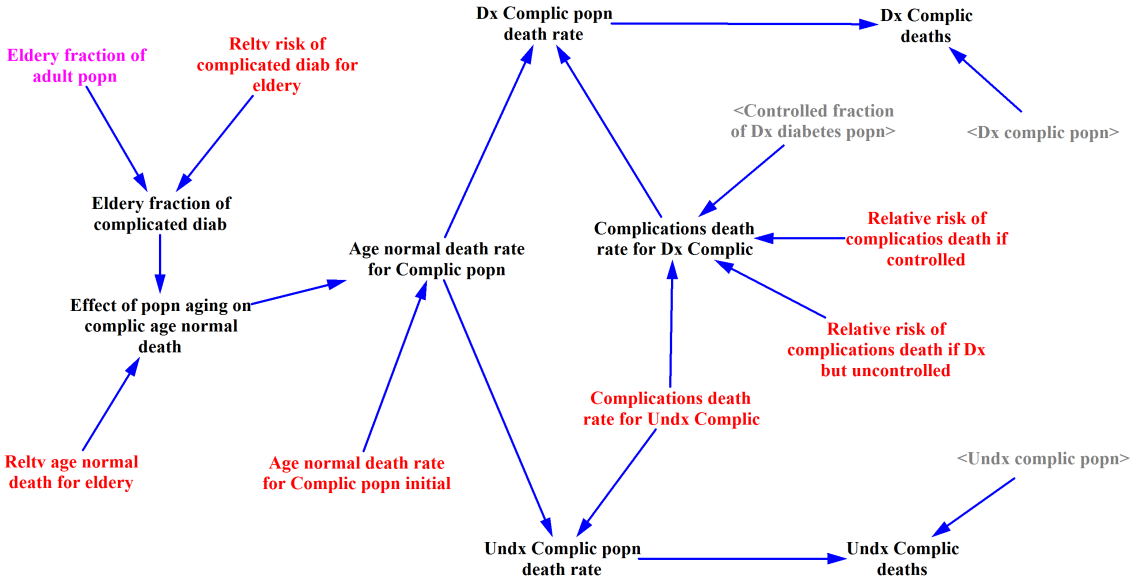

Figure 2.4: The causal structure for the diabetes with complications deaths.

2.3 Disease Onset, Recovery and Progression

The onset of prediabetes is affected by two factors, the elderly fraction of the population and the obese fraction of the normoglycemic population, as can be seen in figure 2.5. Obesity is the number one modifiable risk factor in the development of prediabetes. The respective time dependent effects, as multiplicative factors for a

given initial rate, are calculated analogous to equation 2.1. The prediabetes onset rate is then the rate at which the normoglycemic population level flows out into the undetected prediabetes level.

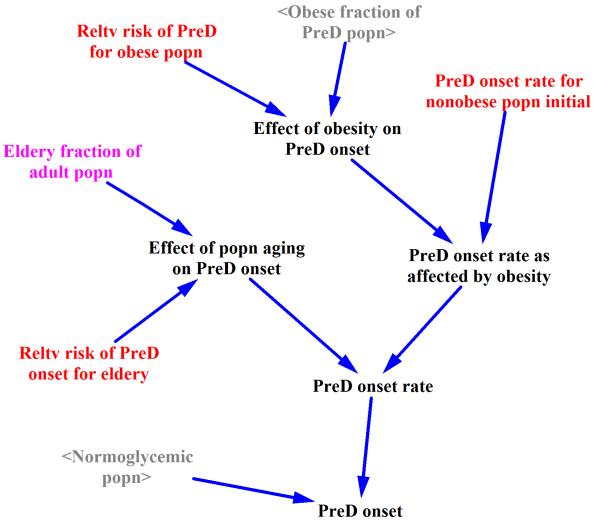

Figure 2.5: The causal structure for the prediabetes onset.

The causal structure for the recovery from prediabetes is seen in figure 2.6. It is assumed that a constant percentage of the prediabetes population heals by itself every year. This stands for the natural fluctuation exchange between the normoglycemic and the prediabetes population levels. The other healing is due to weight loss. That these are the only two relevant influencing factors is supported by three arguments:

- Firstly obesity is the main modifiable risk factor.
- Secondly smaller factors are taken into account aggregated in the constant healing rate.
- And thirdly the contribution of drugs for the treatment of prediabetes, although they are just clinically tested, is negligible.

Additionally the favored choices for the treatment of prediabetes are more physical activity, a healthier diet and weight control due to the other positive and desirable

effects they have compared to a treatment with medicaments. The equation for the prediabetes recovery rate is

> PreD recovery rate = PreD recovery rate normal+Fraction of obesity reduction resulting in PreD recovery*IF THEN ELSE(Rate of change in PreD obese fraction¡0, ABS(Rate of change in PreD obese fraction)*Obese fraction of PreD popn, 0) ,

incorporating the assumption that an additional prediabetes recovery is only obtained if the fraction of obese people in the prediabetes population level decreases.

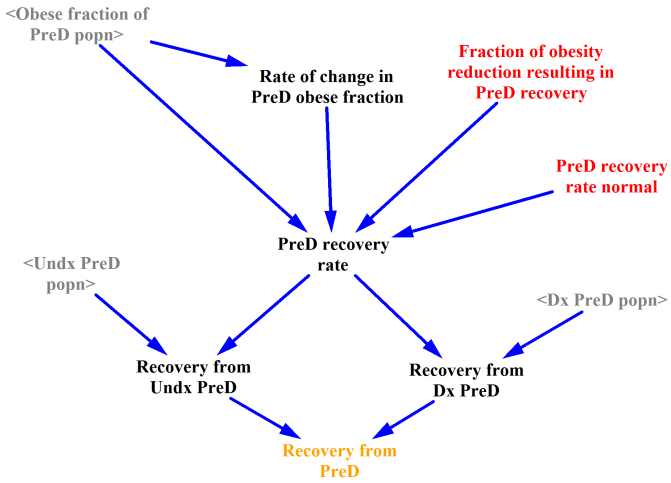

Figure 2.6: The causal structure for the prediabetes recovery.

The causal structure for the onset of diabetes from prediabetes can be seen in figure 2.7. Since type-2 diabetes does not occur instantaneously but is developed gradually over time, prediabetes is a necessary precursor of diabetes. The onset takes always place unobserved, since the blood glucose levels are, even if prediabetes is detected, not measured on a day to day basis. The onset rate is again affected by the elderly fraction of the adult population and the obese fraction of the prediabetes population. It is calculated in the same way as for the prediabetes onset. The difference here are

not only the absolute values of the input parameters, but also that the progression rate for the detected prediabetes population is lowered if it is controlled:

> Diabetes onset rate for Dx PreD = Controlled fraction of Dx PreD popn*Relative risk of diabetes onset if controlled PreD+(1-Controlled fraction of Dx PreD popn)*Diabetes onset rate for uncontrolled PreD

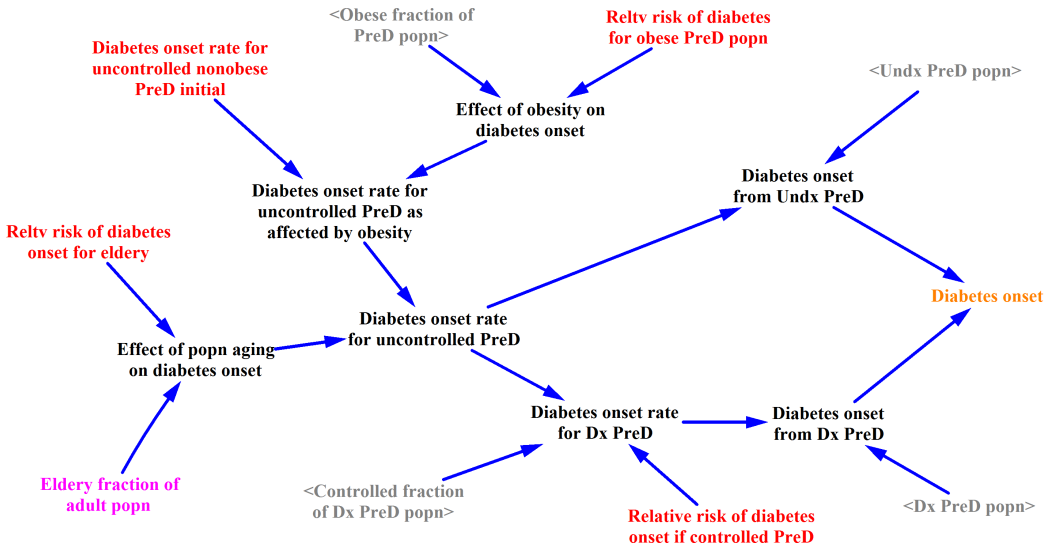

Figure 2.7: The causal structure for the diabetes onset.

The respective flows are, as always, calculated as the product of the levels times the onset rates. The total number of people developing diabetes per year, the yellow Diabetes onset, is one of the output variables which is of interest for health care strategies. The main goal of the simulation system is to find a policy with which it is possible to keep this flow low, since there is no healing of type-2 diabetes.
The logical structure of the calculation of progression rates from diabetes to diabetes with complications can be seen in figure 2.8. If diabetes is undetected, then there is a constant probability for the development of complications per year. If diabetes is detected and controlled this rate can be lowered:

> Progression rate for Dx Uncomp = Progression rate for Undx Uncomp*(Controlled fraction of Dx diabetes popn*Relative risk of progression if controlled+(1-Controlled fraction of Dx diabetes popn)*Reltv risk of progression if Dx but uncontrolled)

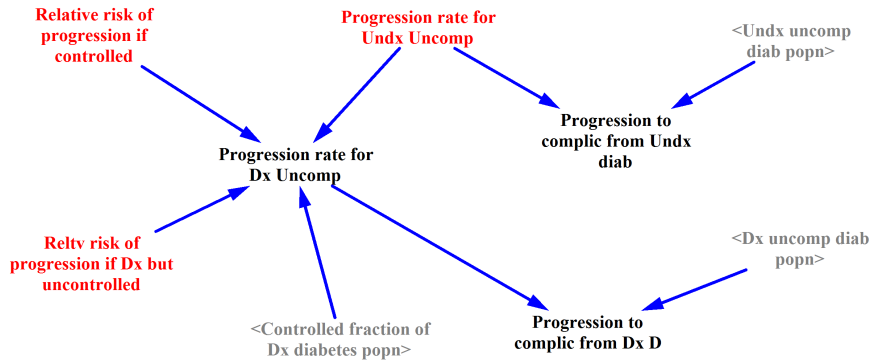

Figure 2.8: The causal structure for the progression from diabetes to diabetes with complications.

These rates are neither affected by the obese nor by the elderly part of the population. The obese fraction plays no role since the damage to blood vessels and nerves is due to a high blood sugar level and not due to adipositas. That the elderly part of the population plays no role has two reasons. The first is that if diabetes is undetected then complications occur after some time independently of the age. The second is that we are interested only in the complications caused by diabetes and not caused in the same diseases as a consequence of age. This argument is also valid for leaving out the obese fraction of the population.

2.4 Diagnosis

The diagnosis flows from the undetected to the detected population levels in the stocks and flows diagram on page 16 are more complicated to calculate. The detection rates depend on the progression rates. They are also influenced by the average time between medical doctor visits, the significance level off the tests used (c. f. section 1.1) and the fraction of people seeking health care. They are delayed, since the progression usually occurs unobserved. All delays are always first order exponential delays, but changing them, for example to a power law delay, does not alter the qualitative behavior. Also, if always a constant percentage would be diagnosed,

there would be a finite limit value of undiagnosed cases, which could be calculated as the sum of a geometric series. The corresponding growth behavior then tuns out to be an exponential one. There are always more ways a diagnosis can be obtained, depending on whether the disease progression happened recently or the people are already in the undiagnosed level for some time. The details for the three diagnosis flows are given in the following.

First there is the diagnosis of prediabetes, for which the causal structure can be seen in figure 2.9. All the gray variables in brackets were already defined in previous diagrams. The prediabetes diagnosis flow is calculated as

PreD diagnosis =
DELAY1(PreD onset,Avg time from onset to detection for Undx hyperglycemic)*(1-(Death rate for nondiab popn+Diabetes onset rate for uncontrolled PreD)*Avg time from onset to detection for Undx hyperglycemic)*Dx fraction of recent PreD onset
+
(Undx PreD popn/Avg time from onset to detection for Undx hyperglycemic)*(Dx fraction of recent PreD onset-DELAY1(Dx fraction of recent PreD onset,Avg time from onset to detection for Undx hyperglycemic))/(1-DELAY1(Dx fraction of recent PreD onset,Avg time from onset to detection for Undx hyperglycemic)) .

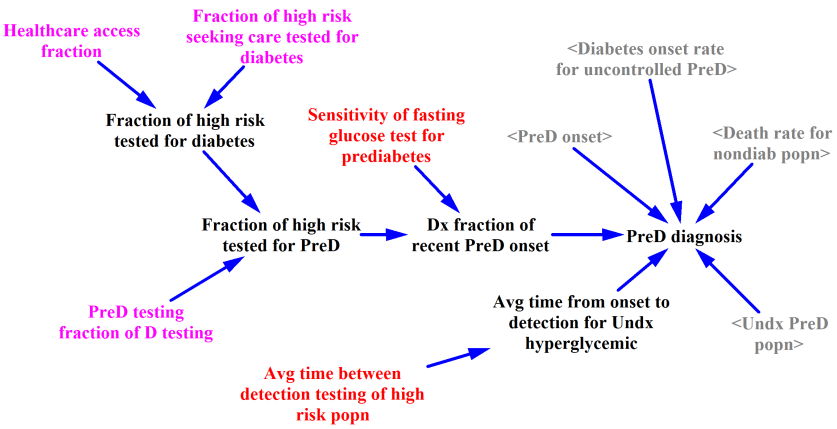

Figure 2.9: The causal structure for the diagnosis of prediabetes.

The two parts of the equation, separated by the plus sign, represent two particular ways to the diagnosis of prediabetes. The first part comes from the diagnoses of people with recent-onset prediabetes. The total prediabetes onsets are diagnosed with a first order exponential time delay, where the time constant is given by half of the average time interval between two medical doctor visits which are calculated in Austria from the total number of doctor visits. They have an average value of 2 to 3 years. Next this rate has to be reduced by the percentage of people which died or progressed to diabetes in the average time from onset to diagnosis. Then the rate is multiplied by the diagnosis fraction for recent-onset prediabetes cases, which is the product of the fractions of high-risks screened for prediabetes times the sensitivity of the test used (usually a FPGT, since it is more sensitive for prediabetes than the OGTT).
The fraction of high risks tested for prediabetes is a fraction of the high-risks tested for diabetes with just another multiplicative conjunction. The second part of the prediabetes diagnosis equation is due to the people whose prediabetes was not diagnosed soon after onset. It is calculated as the yearly part of the undetected prediabetes level multiplied by a ratio. The ratio is the change of the detection fraction of recent-onset prediabetes compared to the last period through the fraction of people who did not get diagnosed in the past period.
The logical structure for the calculation of the diabetes diagnosis flow in figure 2.10 is more complicated. The sensitivity of testing for uncomplicated diabetes is just a statistical mixture of the sensitivities of the FPGT and the OGTT, depending on how often they are used. The fraction of the high risk persons tested for diabetes is the product of the high risks seeking care and the health care access fraction. In Austria the health care access fraction is one, due to the obligatory health insurance, so this factor could be omitted. This is one of the main differences comparing with the original system discribed by J. Homer et al. [22].
Nonetheless we decided to keep these input parameter because it can be used for the simulation of socially disadvantaged groups. The product of the tested fraction and the sensitivity gives the detected fraction of the population with a recent onset of diabetes. The uncomplicated diabetes diagnosis flow is calculated as the sum of three contributing factors (which are again separated by a plus sign):

Uncomplicated diab diagnosis =
DELAY1(Diabetes onset from Dx PreD,Avg time from diab onset to MD visit for

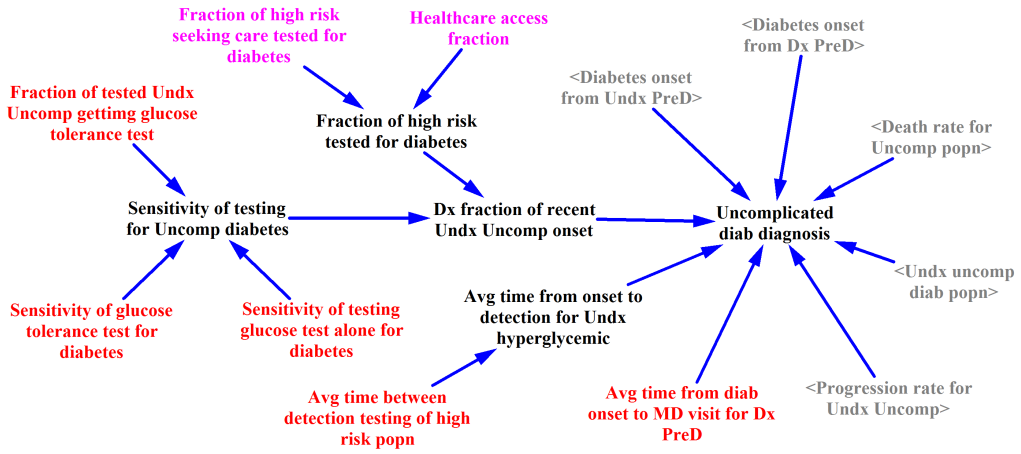

Figure 2.10: The causal structure for the diagnosis of diabetes.

Dx PreD)

+

DELAY1(Diabetes onset from Undx PreD,Avg time from onset to detection for Undx hyperglycemic)*(1-(Death rate for Uncomp popn+Progression rate for Undx Uncomp)*Avg time from onset to detection for Undx hyperglycemic)*Dx fraction of recent Undx Uncomp onset

+

(Undx uncomp diab popn/Avg time from onset to detection for Undx hyperglycemic)*(Dx fraction of recent Undx Uncomp onset-DELAY1(Dx fraction of recent Undx Uncomp onset,Avg time from onset to detection for Undx hyperglycemic))/(1-DELAY1(Dx fraction of recent Undx Uncomp onset,Avg time from onset to detection for Undx hyperglycemic))

The first part accrues from the detected prediabetes population level. It is assumed that these people visit a doctor regularly, so diabetes onset is diagnosed in all such cases after some delay. The second part of the equation accounts for those people who progressed from the undetected prediabetes level to the undetected diabetes level recently. This is done in the same manner as in the prediabetes case. This is also true for the third part, which describes the diabetes detection for people who live already a longer time with undetected diabetes.

The causal structure for the diagnosis of complicated diabetes is presented in figure

2.11. The testing sensitivity is again the average sensitivity of the tests used. The fraction of the complicated population seeking health care is the sum of fraction who have ready access to health care and, for those who do not, the fraction of people with symptoms bad enough that they show up at a health care institution anyway. This last block can be left away in Austria. The detection fraction of the people with a recent onset of the complications is the product of the three preceding variables. The equation for the calculation of the diagnosis flow for diabetes with complications

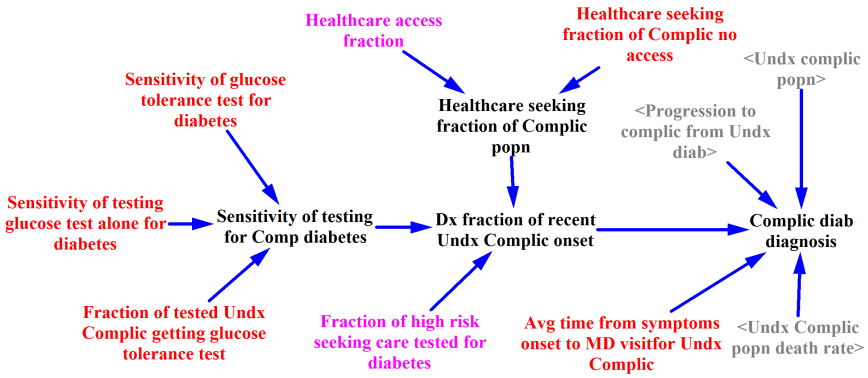

Figure 2.11: The causal structure for the diagnosis of diabetes with complications.

has the same logic as for the prediabetes diagnosis case and is given by:

Complic diab diagnosis =
DELAY1(Progression to complic from Undx diab,Avg time from symptoms onset to MD visit for Undx Complic)*(1-Undx Complic popn death rate *Avg time from symptoms onset to MD visit for Undx Complic)*Dx fraction of recent Undx Complic onset
+
(Undx complic popn/Avg time from symptoms onset to MD visit for Undx Complic)*(Dx fraction of recent Undx Complic onset-DELAY1(Dx fraction of recent Undx Complic onset,Avg time from symptoms onset to MD visit for Undx Complic))/(1-DELAY1(Dx fraction of recent Undx Complic onset,Avg time from symptoms onset to MD visit for Undx Complic))

2.5 Obesity

In this model obesity influences the onset of prediabetes (figure 2.5), the recovery of prediabetes (figure 2.6) and the onset of diabetes (figure 2.7). Obesity is modeled by the average weight and through it the average body mass index (BMI). The BMI is the weight in kilograms of a person divided by the height in meters to the square:

$$\text{BMI} = \frac{\text{weight}}{\text{height}^2} \quad [\text{kg/m}^2] \tag{2.2}$$

There are several indices for the ratio between the body height and bodyweight around, but the BMI proved to be the best for all age categories. It accounts better for the different heights and there is a strong correlation to the fat fraction of body weight. The later is especially important, since it is responsible for the higher health risks associated with overweight. An BMI from 20 to 25 kg/sqm. is considered normal, below 20 kg/sqm. people are under wighted and above 25 kg/sqm. they are overweighted. An BMI over 30 is considered obese and diagnosed as adipostity.
The average BMI of the adult population is calculated from historical height data and the average body weight. The body weight is modeled as a level with a two-way flow which accumulates changes over time, as can be seen in figure 2.12. Note that this part constitutes a separate model. It enters the rest of the model just via the obese fraction of the adult population, which is an empirical lookup function of the average BMI. The original functions from [25] are:

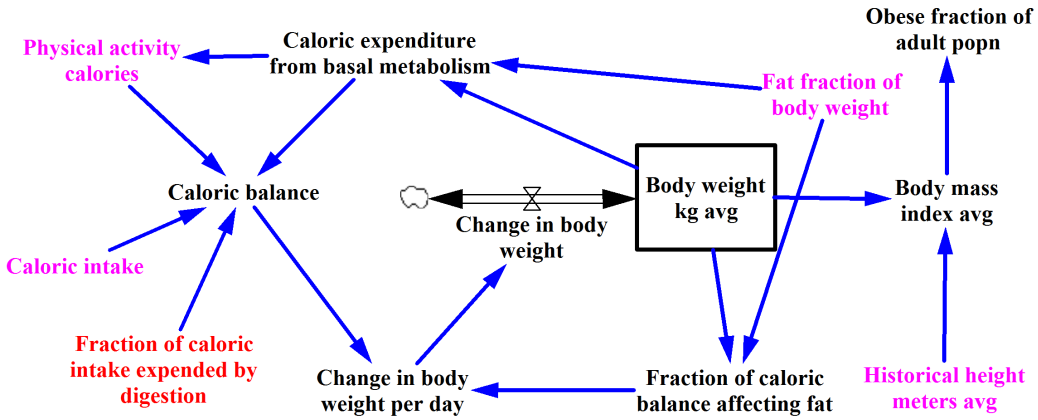

Figure 2.12: The caloric balance feedback loop and the causal structure leading to the obese fraction of the adult population.

Change in body weight per day =
Caloric balance*(Fraction of caloric balance affecting fat/9300+(1-Fraction of caloric balance affecting fat)/4100)

where

Fraction of caloric balance affecting fat = 1/(1+((10.4/(Body weight kg avg*Fat fraction of body weight))*(4100/9300)))
Caloric expenditure from basal metabolism = (Body weight kg avg*(0.024*Fat fraction of body weight+0.102*(1-Fat fraction of body weight))+0.85)*238.7
Caloric balance = Caloric intake-Physical activity calories-Caloric expenditure from basal metabolism-(Caloric intake*Fraction of caloric intake expended by digestion)

These equations need to be adopted, since not all the required historical input data is available in Austria. The new equation for the caloric balance is altered so that the calories spent by physical activity is given as a multiple of the basal metabolic rate. This ratio is called the physical activity level and varies between 1.2 and 2.4. So the new equation is

Caloric balance = Caloric intake-Physical activity level*Caloric expenditure from basal metabolism-(Caloric intake*Fraction of caloric intake expended by digestion)

and figure 2.12 should be altered accordingly. The effect is that heavier people doing the same physical exercises for the same time loose more weight than lighter ones. The overall change in the obese fraction of the adult population is a minor one (c. f. section 4.1) but it represents the possible intervention policies better. E. g. the WHO and subsequently the Austrian ministry for health recommend half an hour of exercises per day for the prevention of obesity, type-2 diabetes and vascular diseases. With the obese fraction of the adult population the obese fractions of the normoglycemic, the prediabetes and the diabetes population can now be calculated. The causal structure is seen in figure 2.13. The first striking feature is that the obese fraction of the prediabetes population and of the diabetes population have a back coupling to the obese fraction of the normoglycemic population. This constitutes a problem since these are just calculated variables and not levels. Therefore this logic structure leads to implicit equations. With initial conditions these can be solved self-consistently for the obese fraction of the normoglycemic population and after

an initialization the calculation can then be continued explicitly. Another way to overcome this obstacle would be to introduce accumulating levels between the normoglycemic and the prediabetes fraction as well as between the prediabetes and the diabetes fraction. The second striking feature is that there are many orange variables, which are calculated straightforward from previous results as the respective percentiles. That they are orange, which stands for output variables, signifies that these variables need not necessarily be calculated, but they are quite interesting to observe over the course of time and the equations get a little bit shorter. These

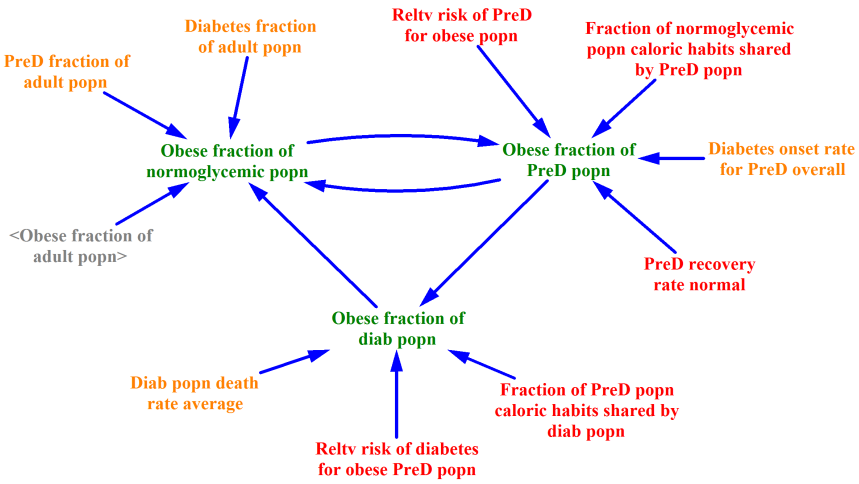

Figure 2.13: The causal structure for the obese fraction of the different levels.

equations for the obese fractions are:

Obese fraction of normoglycemic popn =
(Obese fraction of adult popn-Obese fraction of PreD popn*PreD fraction of adult popn-Obese fraction of diab popn*Diabetes fraction of adult popn)/(1-PreD fraction of adult popn-Diabetes fraction of adult popn)

Obese fraction of PreD popn =
SMOOTHI((Reltv risk of PreD for obese popn*Obese fraction of normoglycemic popn)/(Reltv risk of PreD for obese popn*Obese fraction of normoglycemic popn +(1-Obese fraction of normoglycemic popn)), (1-Fraction of normoglycemic popn caloric habits shared by PreD popn)/(Diabetes onset rate for PreD overall +Death

rate for nondiab popn+PreD recovery rate normal) , Obese fraction of PreD popn initial)

Obese fraction of diab popn =
SMOOTHI((Reltv risk of diabetes for obese PreD popn*Obese fraction of PreD popn)/(Reltv risk of diabetes for obese PreD popn*Obese fraction of PreD popn+(1-Obese fraction of PreD popn)), (1-Fraction of PreD popn caloric habits shared by diab popn)/Diab popn death rate average, Obese fraction of diab popn initial)

The calculation of the obese fraction of the normoglycemic population is done algebraically from the other fractions. It is just the respective percentile of the adult population and, as said before, the initial percentile is calculated implicitly from the other two fractions and their initial values. The obese fraction of the prediabetes population is calculated as a first order exponential delay of the obese fraction of the normoglycemic population. In the steady state it is assumed to be a multiple risk factor higher than for the normoglycemic population. The smoothing results form the ,in the average, elder prediabetes population compared to the normoglycemic one and from different nutrition and physical activity habits in the different levels. The delay time constant can be viewed as the average time people spent in the prediabetes population levels, based on the rates of outflow due to death, prediabetes recovery and diabetes onset. The obese fraction of the diabetes population is calculated in the same way.

2.6 Disease Management

Disease management programs are a global system approach to the health care challenges of epedemics [26], chronic diseases [35] and diseases of civilization. For the system at hand only the parts concerning the patients and the development of diabetes are of interest. The influencing factors and their logical connections taken into account in this system can be seen in figure 2.14. Generally speaking this part of the system describes the fractions of the prediabetes and the diabetes populations which have their blood glucose level under control.
The controlled fractions are the product of the managed fraction of the detected population and the respective controlled fractions of this managed population. The equations for the controlled fractions of the managed population are

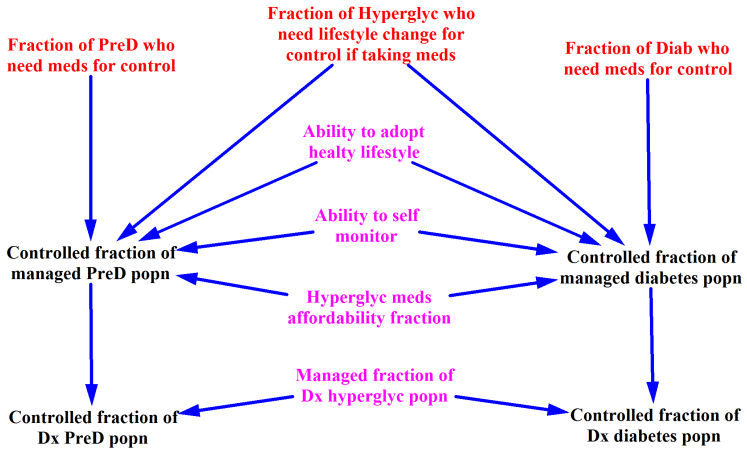

Figure 2.14: The causal structure for the controlled fractions of the people with diagnosed prediabetes and diabetes.

Controlled fraction of managed PreD popn =
(1-Fraction of PreD who need meds for control)*Ability to adopt healty lifestyle + Fraction of PreD who need meds for control*Ability to self monitor*Hyperglyc meds affordability fraction*(1-Fraction of Hyperglyc who need lifestyle change for control if taking meds+Fraction of Hyperglyc who need lifestyle change for control if taking meds*Ability to adopt healty lifestyle)
Controlled fraction of managed diabetes popn =
(1-Fraction of Diab who need meds for control)*Ability to adopt healty lifestyle + Fraction of Diab who need meds for control*Ability to self monitor*Hyperglyc meds affordability fraction*(1-Fraction of Hyperglyc who need lifestyle change for control if taking meds+Fraction of Hyperglyc who need lifestyle change for control if taking meds*Ability to adopt healty lifestyle) .

The first part of the equations represents the people who, in each case, do not need medicaments for control. The second part describes the people who need drugs. These controlled fractions are influenced by the ability to self-monitor to assess the effectiveness of the drugs and by the affordability, which is in Austria always one, except maybe for socially disadvantaged groups. A third influencing factor is the ability to adopt a healthy lifestyle if drugs alone are not enough to control the blood glucose level.

2.7 Health Care Costs

The costs of type-2 diabetes are quite differently from the USA. On the one hand there are the acute costs, which include hospital inpatient stays, emergency room visits and ambulance services. On the other hand there are the continual costs including physician office visits, medications, glucose testing and preventive measures. All costs are calculated as a total per year based on values from 2005/2006 without inflation.

The causal structure for the calculation of the continual health care costs can be seen in figure 2.15. These are always the initial values times the current fraction of the affected people. Note that all costs are just output variables and they have no feedback effect on the rest of the system (although they are directly influenced by some input variables). The values in Euro could be corrected easily for macroscopic effects like inflation or the growth of the gross domestic product. The growth rates remain the same in any case.

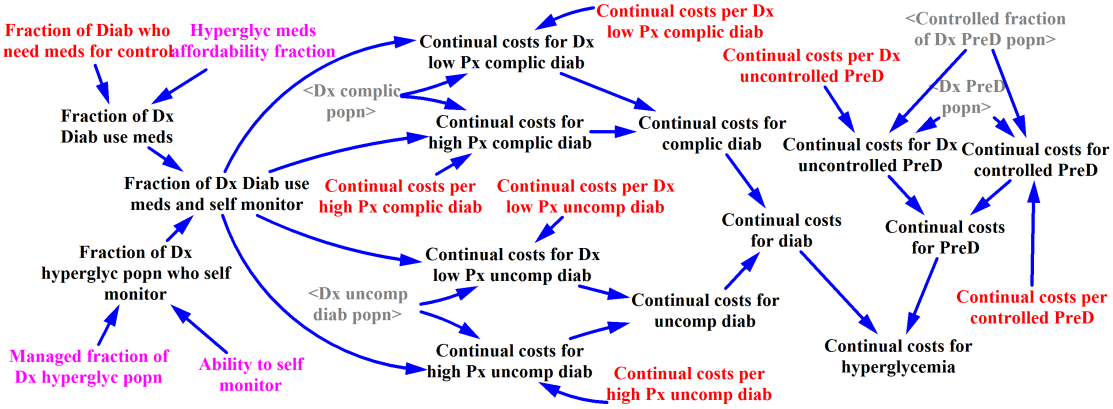

Figure 2.15: The causal structure for the continual health care costs.

Figure 2.16 shows the causal structure for calculating the acute and the total health care costs. They are calculated as the costs per person times the number of affected people and respectively as the sum of the individual contributions.

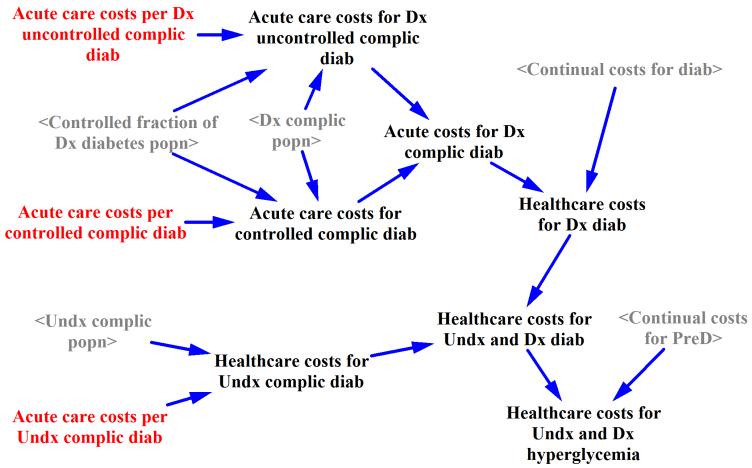

Figure 2.16: The causal structure for the acute health care costs.

Chapter 3

Input Data

One major problem when building this model was finding the correct input data. In this chapter the encountered problems will be discussed generally and then a list of the time dependent and the time independent input parameters in order of their appearance will be given.

The aim of this SD model is to compare different policies in complex systems with many interactions and nonlinearities. These systems may produce counterintuitive behavior. In order to built a realistic model as many variables as possible, which might have a direct or an indirect influence, have to be incorporated. This is in conflict with the aim to make a manageable model. Although in the original model no distinction with respect to sex was included we did split the calculation for men and women. The reason is that the risk factors are very different and therefore other interventions may be necessary. So the split was not done due to causalities but with a different aim in mind.

Not all variables used in the model, although needed in the logical structure, are measured or available, e. g. the undiagnosed population levels. Nonetheless these variables are needed because one is interested in how a change in these variables can influence the output data. When no direct data was available the parameters were estimated. Sometimes there were indirect indicators and at other times they had to be found experimentally. This was done by tuning the uncertain variables manually but applying them to an as small as possible part of the model so that the available historical data were correctly reproduced. This is called the partial-model estimation approach.

Another problem encountered is that the available data originates from different

years, for example mikrocensus data in more or less regular intervals. In the model this was circumvented, where applicable, by the use of time-constant risk factors (c. f. equation 2.1 on page 17).

Another problem is that the available data has a very different accuracy and statistical variation. To formulate the solution shortly: we just expected the worst and analyzed the stability with the same relative magnitude of error for all variables (c. f. section 4.1).

When no data for Austria was available we used data from similar western countries, like Bavaria in Germany or even the USA. In the following we give the values of all variables used and reference their origin.

3.1 Time Independent

We give here tables with the time-independent input parameters in order of their appearance in the model. These are those parameters displayed in red in the previous chapter. The initial values are from the year 1980, while time-constant odds may come from different years. Actually if those parameters could be calculated from available data originating from different years, we used this to verify that they really stay constant. In the first column the name of the parameter is listed, followed by its value for Austria (Aut) and then the values for the male (m) and female (f) population. In the last column the sources (Src.) of the data are given. Almost all values are derived quantities from the data given in the sources by elementary mathematics. When this is not the case it will be denoted by the letter 'o' for 'original data' in the citation column.

To give an example: the initial death rate for the adult population, the first parameter below, is calculated from references [27] and [28] by taking the death rate per 1000 people and the total number of people in 5-years age categories. With these the total number of deaths of adult people ($\geq$ 20 years) is calculated. By dividing this by the total adult population we get the desired death rate, expressed as a fraction per year. The same is done for the male and the female population.

Name	Aut	m	f	Src.
Death rate for adult popn initial	0.0121	0.0116	0.0126	[27], [28]
Reltv age normal death for elderly	17.51	14.00	24.97	[27], [28]
Death rate for nondiab popn initial	0.0118	0.0097	0,0105	[11]

Table 3.1: The parameters for figure 2.2 on page 19.

Name	Aut	m	f	Src.
Reltv risk of uncomplicated diab for elderly	1.9	1.7	2.0	[11]
Death rate for Uncomp popn initial	0.0124	0.0151	0.0102	[11]

Table 3.2: The parameters for figure 2.3 on page 19.

Name	Aut	m	f	Src.
Reltv risk of complicated diab for elderly	7.6	7.6	7.6	[11]
Age normal death rate for Complic popn initial	0.0282	0.0270	0.0294	[11], [22]
Complications death rate for Undx Complic	0.121	0.116	0.126	[11], [22]
Relative risk of complications death if Dx but uncontrolled	0.465	0.465	0.465	[11], [22]
Relative risk of complications death if controlled	0.167	0.167	0.167	[11], [22]

Table 3.3: The parameters for figure 2.4 on page 20.

Name	Aut	m	f	Src.
Reltv risk of PreD for obese popn	2.6	2.6	2.6	[22]
PreD onset rate for nonobese popn initial	0.043	0.043	0.043	[22]
Reltv risk of PreD onset for elderly	1.15	1.15	1.15	[22]

Table 3.4: The parameters for figure 2.5 on page 21.

Name	Aut	m	f	Src.
PreD recovery rate normal	0.1	0.1	0.1	[22], o
Fraction of obesity reduction resulting in PreD recovery	0.5	0.5	0.5	[22], o

Table 3.5: The parameters for figure 2.6 on page 22.

Name	Aut	m	f	Src.
Diabetes onset rate for uncontrolled nonobese PreD initial	0.0135	0.0142	0.0129	[22], o
Reltv risk of diabetes onset for elderly	1.52	1.44	1.60	[11]
Reltv risk of diabetes for obese PreD popn	2.6	2.6	2.6	[11]
Relative risk of diabetes onset if controlled PreD	0.58	0.58	0.58	[11]

Table 3.6: The parameters for figure 2.7 on page 23.

Name	Aut	m	f	Src.
Relative risk of progression if controlled	0.36	0.36	0.36	[22], o
Progression rate for Undx Uncomp	0.079	0.079	0.079	[11]
Reltv risk of progression if Dx but uncontrolled	1	1	1	[11]

Table 3.7: The parameters for figure 2.8 on page 24.

Name	Aut	m	f	Src.
Sensitivity of fasting glucose test for prediabetes	0.84	0.84	0.84	[22], o
Avg time between detection testing of high risk popn	3	3	3	[11]

Table 3.8: The parameters for figure 2.9 on page 25.

Name	Aut	m	f	Src.
Fraction of tested Undx Uncomp gettimg glucose tolerance test	0.6	0.6	0.6	[11]
Sensitivity of glucose tolerance test for diabetes	0.97	0.97	0.97	[11]
Sensitivity of testing glucose test alone for diabetes	0.84	0.84	0.84	[11]
Avg time between detection testing of high risk popn	3	3	3	[11]
Avg time from diab onset to MD visit for Dx PreD	0.5	0.5	0.5	[11]

Table 3.9: The parameters for figure 2.10 on page 27.

Name	Aut	m	f	Src.
Fraction of tested Undx Complic getting glucose tolerance test	0.9	0.9	0.9	[11]
Healthcare seeking fraction of Complic no access	1	1	1	[11]
Avg time from symptoms onset to MD visitfor Undx Complic	1	1	1	[11]

Table 3.10: The parameters for figure 2.11 on page 28.

Name	Aut	m	f	Src.
Physical activity level	1.3	1.3	1.3	[30], [31]
Fraction of caloric intake expended by digestion	0.1	0.1	0.1	[22]

Table 3.11: The parameters for figure 2.12 on page 29.

Name	Aut	m	f	Src.
Fraction of normoglycemic popn caloric habits shared by PreD popn	0.8	0.8	0.8	[22], o
Fraction of PreD popn caloric habits shared by diab popn	0.2	0.2	0.2	[22], o

Table 3.12: The parameters for figure 2.13 on page 31.

Name	Aut	m	f	Src.
Fraction of PreD who need meds for control	0.33	0.33	0.33	[11], [32], [?]
Fraction of Hyperglyc who need lifestyle change for control if taking meds	0.67	0.67	0.67	[11]
Fraction of Diab who need meds for control	0.95	0.95	0.95	[11], [32], [?]

Table 3.13: The parameters for figure 2.14 on page 33.

Name	Aut	m	f	Src.
Continual costs per high Px complic diab	2150	2150	2150	[32], [?], [22]
Continual costs per Dx low Px complic diab	1050	1050	1050	[32], [?], [22]
Continual costs per high Px uncomp diab	1900	1900	1900	[32], [?], [22]
Continual costs per Dx low Px uncomp diab	765	765	765	[32], [?], [22]
Continual costs per Dx uncontrolled PreD	20	20	20	[32], [?], [22]
Continual costs per controlled PreD	500	500	500	[32], [?], [22]

Table 3.14: The parameters for figure 2.15 on page 34.

Name	Aut	m	f	Src.
Acute care costs per Dx uncontrolled complic diab	8700	8700	8700	[32], [?], [22]
Acute care costs per controlled complic diab	2900	2900	2900	[32], [?], [22]
Acute care costs per Undx complic diab	17400	17400	17400	[32,], [22]

Table 3.15: The parameters for figure 2.16 on page 35.

3.2 Time Dependent

In this section we give the time-dependent input data in order of their first appearance. Graphs of these data between 1980 and 2050 will be given. For the normal scenario it is assumed that all policy dependent values remain constant beginning with 2006.

The first time dependent input variables are the demographic ones: the adult population, that is all people above 20 years of age, and the elderly fraction of the adult population, which is the fraction of the adult population above 65 years. The lower boundary was chosen to be 20 instead of 18 since the available data comes in 5-years age categories. But such a difference does not alter the behavior since almost none of the 18 to 20 year old people develop type-2 diabetes. The upper boundary was chosen to be the legal age of retirement, which usually also includes a significant life style change. If these age borders would be changed the risk factors in the previous section have to be altered accordingly.

In figure 3.1 we see the development of the Austrian adult population and the elderly fraction of it. The middle scenarios of the references [27], [28] and [33] were chosen. The predictions for the future development are usually given in 5 year steps, so we need to interpolate them. Note that the behavior of the systems does not change significantly if the available data points are interpolated linearly or by 3^{rd} order Splines. In the next two figures (3.2 and 3.3) the respective graphs for the male and the female population are given. We see that there are more females and that they get, in the average, older than their male counterparts. Both elderly fractions are increasing, corresponding to the aging of the society, but the male fraction is increasing faster than the female one.

The following time dependent input variables are all given between 1980 and 2005 and are expected to stay constant after that time in the basis scenario. This is adequate because we get a result that describes what would happen if the status quo would be prolonged. So this is a plausible reference scenario.

In the original model there is the fraction of people who have an access to health care services (the Healthcare access fraction in figure 2.9 on page 25). Due to the social security system in Austria this fraction is always set to one (and could therefore be eliminated from the model). However, it may be interesting to study socially disadvantaged groups in the future, where it is known that they for example visit

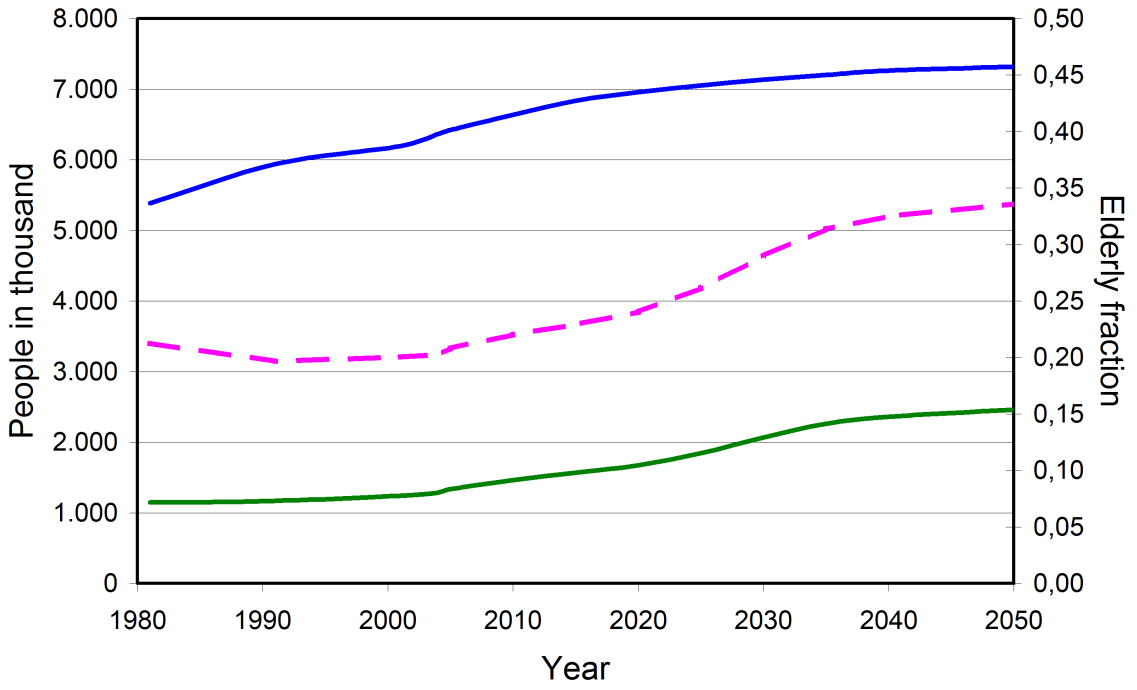

Figure 3.1: The adult (blue) and the elderly (green) population and the fraction of elderly people (rose slashed, right scale) in Austria between 1980 and 2050.

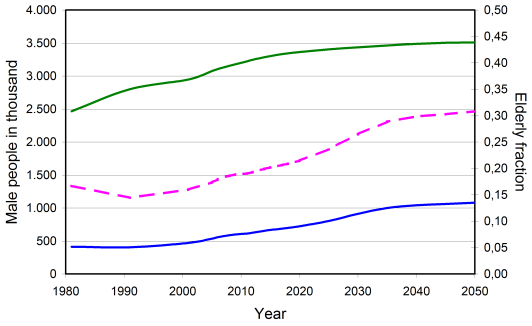

Figure 3.2: The same as figure 3.1, but only for the male part of the population.

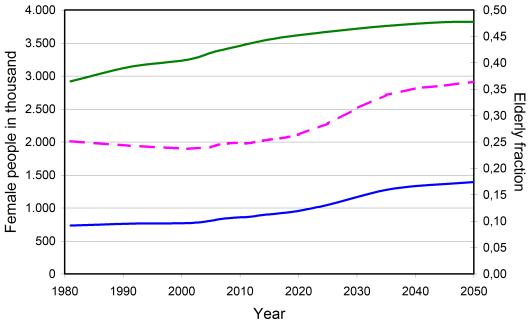

Figure 3.3: The same as figure 3.1, but only for the female part of the population.

doctors 20 percent less than the average population [34].
The next quantity, the Fraction of high risk seeking care tested for diabetes , is rather difficult to obtain over the time. It includes all people where a diabetes screening would be indicated as well as people who exhibit typical symptoms of diabetes complications. A first approximation is the number of diabetes cases diagnosed at general health screenings, which are in Austria free of charge once a year. Now the newly diagnosed cases at hospitals and general practitioners are added. This gives an order of magnitude but systematically underestimates the true number of tested persons, since the negative test results are not included. Also no data on the true data of the high risk persons is available, so the fraction getting tested can only be estimated.
This was done so that the closest possible match of the fraction of diagnosed diabetes cases with available historical data has been obtained. Experts agree that it has risen in the past years, which is also due to a higher public awareness and several education programs [35]. The chosen time-line can be seen in figure 3.4.

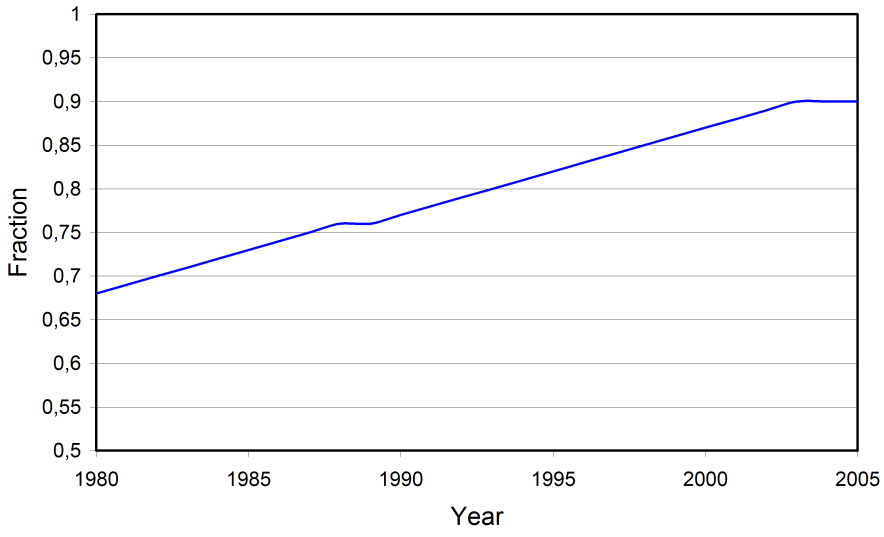

Figure 3.4: The fraction of high risk health care users tested for diabetes between 1980 and 2005.

The testing for prediabetes has only started about a decade ago [11]. In the model the fraction of high risk people screened for prediabetes is expressed as a fraction of the high risk people tested for uncomplicated diabetes described above. This PreD testing fraction of D testing is displayed in figure 3.5. It has again been chosen to give the right historically known values for diagnosed prediabetes.

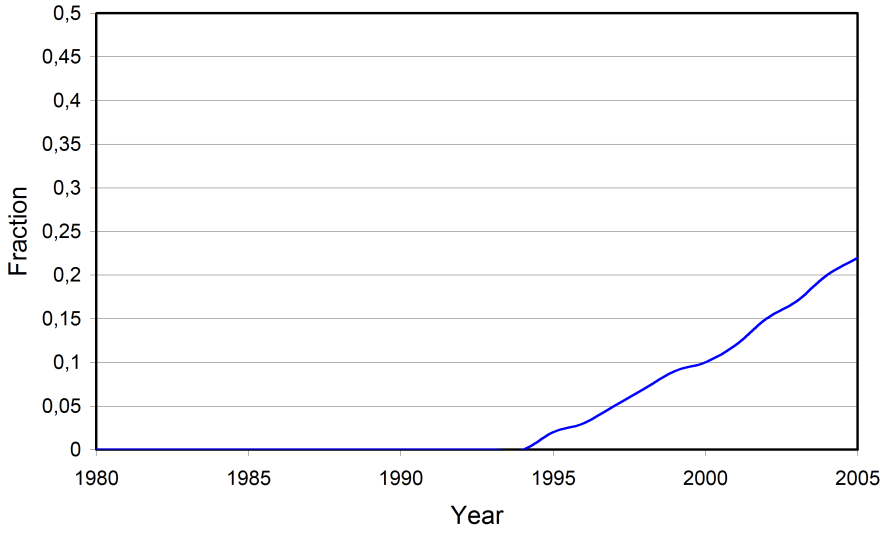

Figure 3.5: The fraction of high risk people tested for prediabetes as a fraction of the high risk people tested for diabetes between 1980 and 2005.

The next four time dependent input variables influence the ability to control prediabetes and diabetes, as seen in figure 2.14 on page 33. It is again fortunate to life in Austria, where the medicaments a person needs are supplied. Therefore the Hyperglyc meds affordability fraction is set to one with the same possibilities as for the health care access fraction concerning marginal groups.
In the formula given in the last chapter it is multiplied by the ability to self monitor. So this is a good example of a variable mathematically not necessary, since only the product of it with another variable enters the system, but which has been introduced in the model to represent a possible way of influencing it. There may even be other, mathematically equivalent ways which produce the same effect but are not (or not as easily) accessible in the real world.
The estimated fraction of detected hyperglycemics who are under regular surveillance can be seen in figure 3.6. They are assumed to be able to have their blood glucose level under control. The data is again obtained through the stabilization method from data out of [11], [3] and [36].
The time development of the Ability to self monitor is given in figure 3.7. It is given as a fraction of the people under management and rises steeply in the past years. This is due to technical advances so that daily blood glucose testing can be done now very simply at home. It is adjusted starting from the original data given in [22].

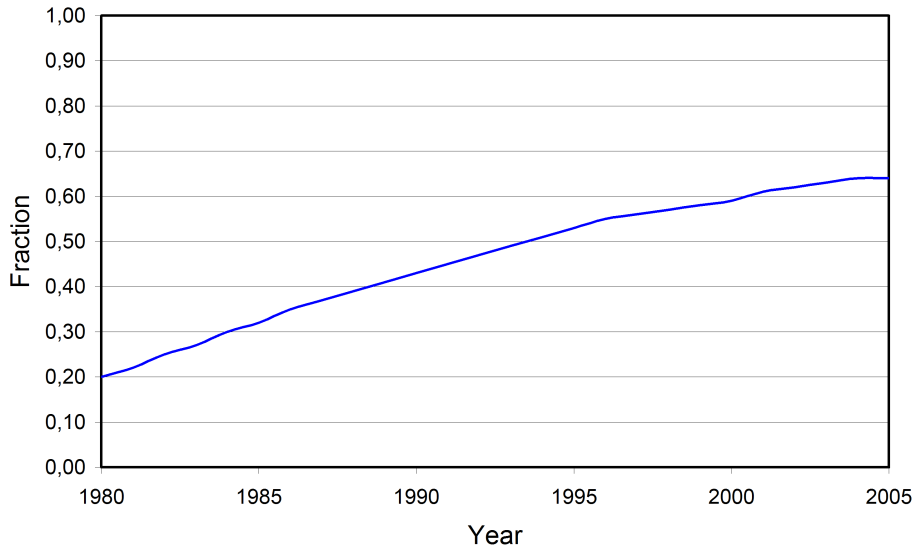

Figure 3.6: The managed fraction of the detected hyperglycemic population between 1980 and 2005.

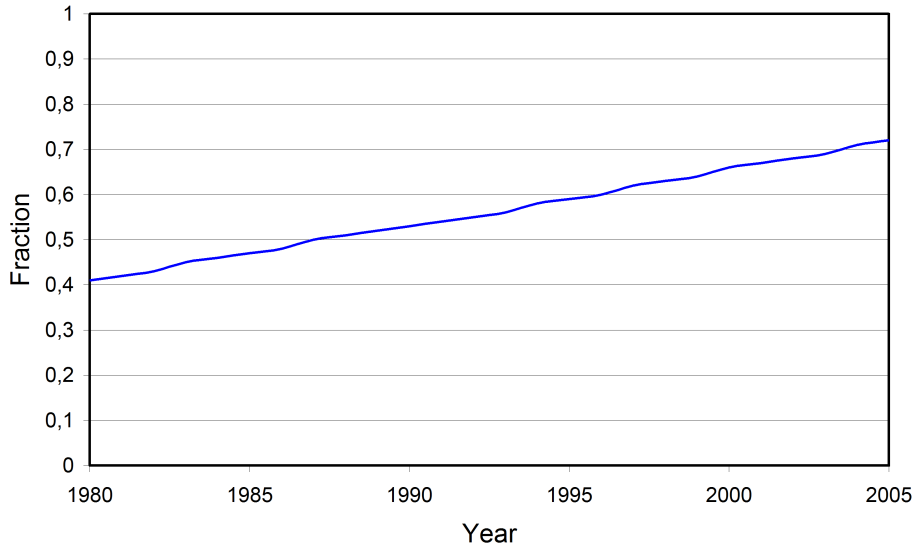

Figure 3.7: The ability to self monitor the blood glucose level between 1980 and 2005.

The last time dependent variable in this block is the Ability to adopt a healthy lifestyle. This includes nutrition and exercises. Parts of this data from different years is available in [30] and [37]. This was then interpolated for the single years and can be seen in figure 3.8. We see that there is still potential left here.

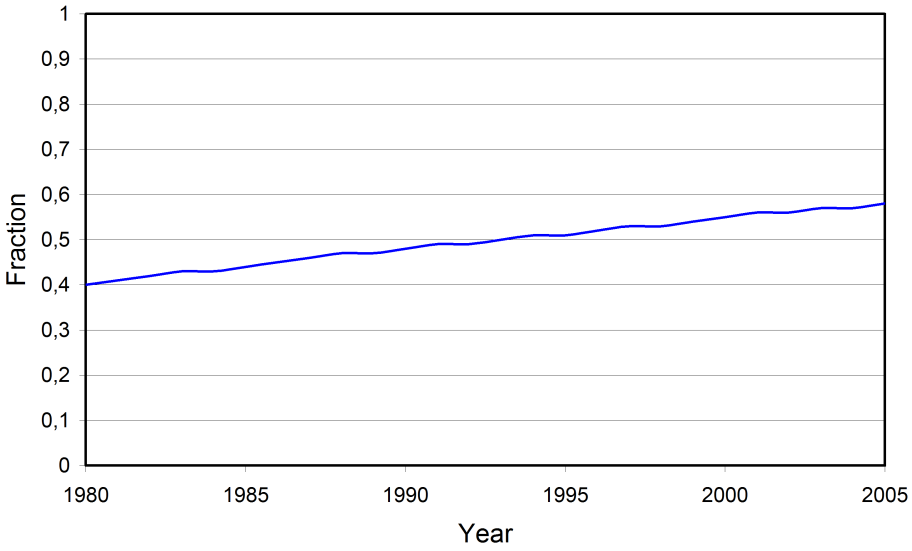

Figure 3.8: The ability to adopt a healthy lifestyle between 1980 and 2005.

The last of the time dependent input variables are connected with the BMI-feedback loop given in figure 2.12 on page 29. The Historical height meters avg is, although in principle time dependent, assumed to be constant.
This has to do with two observations: firstly the time scale over which the average height changes is about 50 years, the average height in the USA changed only about half a centimeter between 1980 and 2001. People tend to get taller so the average height increases. But people are also getting older and older and start shrinking again, so that the average over the whole population increases very slowly. A more detailed SD model for the average BMI or obesity in general would have to include the age dependent nutrition and body height effects. Secondly there are no time series for the average height in Austria. The data used is $175,5$cm for men and $163,8$cm for women from reference [38] or rather [39] out of the year 1991. The gender standardized value is $169,3$cm. Also the Fat fraction of body weight does not change very much, even in the original model is was assumed to be constant 30% [22].
What remains are the Caloric intake and the Physical activity level (instead of the

Physical activity calories). These can be reconstructed out of the historical BMI and nutrition calories data given in [29], [30], [31] and [37]. They are displayed in the figures 3.9 and 3.10.

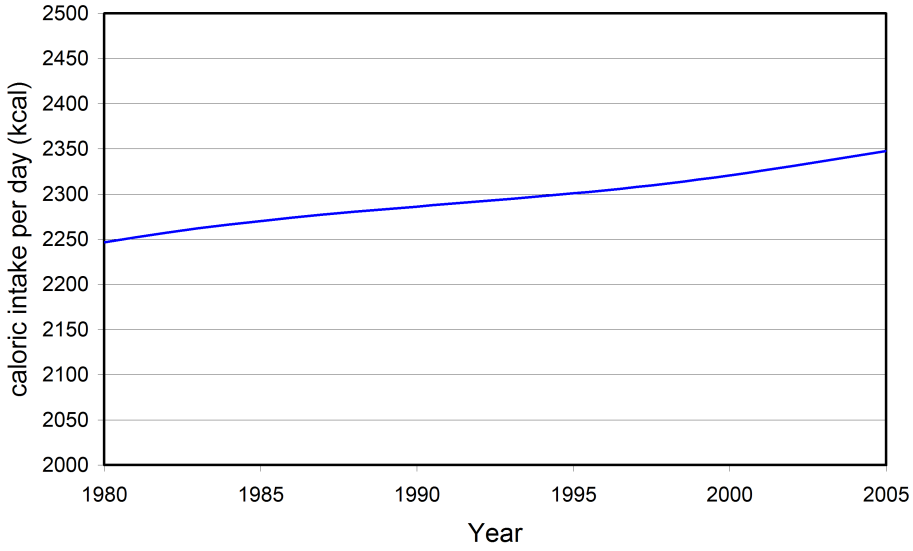

Figure 3.9: The average caloric intake per day between 1980 and 2005.

The output of this feedback loop is the average BMI. The obese fraction of the adult population of is a well known function empirical of the average BMI, $of : \overline{BMI} \rightarrow of(\overline{BMI})$. It is defined as a lookup function, that is its value is looked up in a table. The graph of the function is given in figure 3.11.

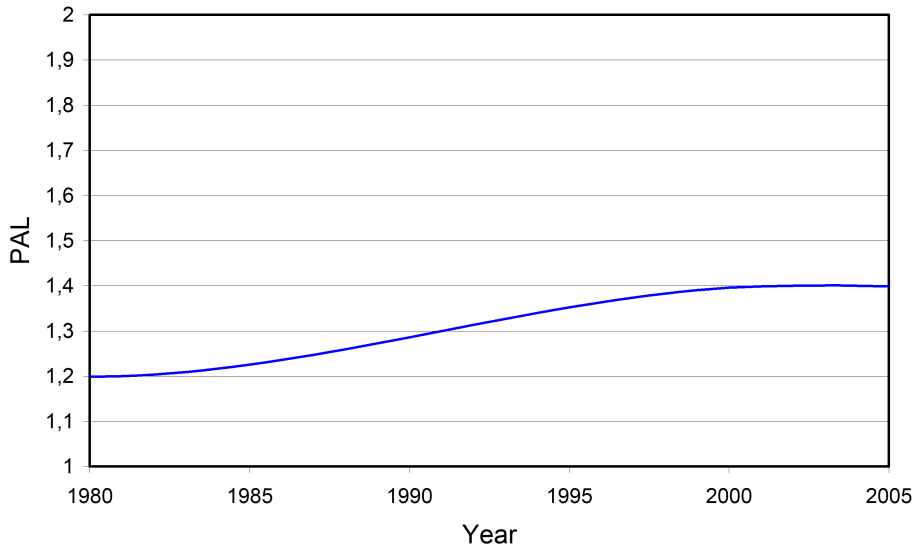

Figure 3.10: The physical activity level between 1980 and 2005.

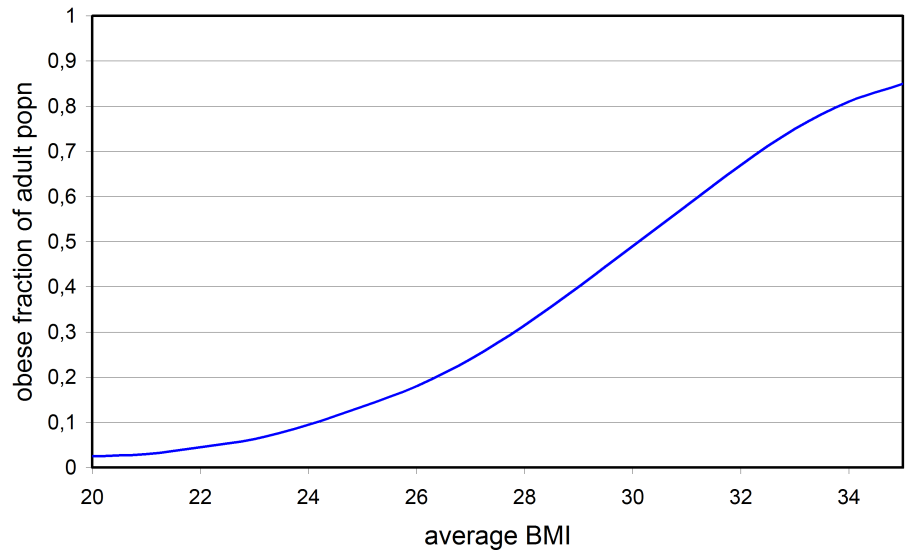

Figure 3.11: The obese fraction of the adult population as a function of the average BMI.

Chapter 4

Output Data

After all the necessary input data is incorporated in the model experiments can be done with it. There are two kinds of interesting output quantities of this SD model of the type-2 diabetes mellitus prevalence in Austria:

- Variables to validate the usefulness of the model and
- variables to make predictions for the future development of the real life system.

From this follows the structure of this chapter: to prove the usefulness of the model two things have to be shown, on the one hand that the model is stable and on the other hand that historical data is reproduceable. After that is done several simulation runs are made and interpreted.

4.1 Stability Analysis

Qualitatively this type-2 diabetes mellitus System Dynamics model is very complex. To show this we give the number of feedback loops for the different population levels (c. f. figure 2.1 on page 16) in table 4.1:
The number of the feedback loops gives the number of different ways one can go through the system starting and ending at the same level. The maximal length of the loops gives the maximal number of other variables visited on such a way. For example the loop number 1222 of length 24 of the Dx uncomp diab popn is:

> **Dx uncomp diab popn** → Progression to complic from Dx D → Dx complic popn → Dx Complic deaths → Diab deaths → Diab popn death rate average →

Level	Number	max. length
Normoglycemic popn	1618	23
Undx PreD popn	3609	25
Dx PreD popn	3025	25
Undx uncomp diab popn	3635	25
Dx uncomp diab popn	1222	24
Undx complic popn	1820	25
Dx complic popn	1831	25

Table 4.1: The different population levels with the number of feedback loops and the maximal length of the loops.

> Obese fraction of diab popn → Obese fraction of normoglycemic popn → Effect of obesity on PreD onset → PreD onset rate as affected by obesity → PreD onset rate → PreD onset → Undx PreD popn → PreD diagnosis → Dx PreD popn → PreD popn → Diabetes onset rate for PreD overall → Obese fraction of PreD popn → Effect of obesity on diabetes onset → Diabetes onset rate for uncontrolled PreD as affected by obesity → Diabetes onset rate for uncontrolled PreD → Diabetes onset rate for Dx PreD → Diabetes onset from Dx PreD → Undx uncomp diab popn → Uncomplicated diab diagnosis → **Dx uncomp diab popn**

We see that the system is very complex and therefore there is no chance to analyze the stability of the system analytically. However, there are some qualitative considerations which can be made:

The first has to do with the stocks and flows structure: the only inflow in the system is the adult population. Nowhere else in the system new population can enter, the people are only allowed to die. Therefore the number of people in the system is bounded and no unlimited growth can arise. Secondly all the flows are time-delayed, so we can expect that the system responds smoothly to discrete changes of the input variables and no instantaneous depletion of a level can occur.

And finally an first order analytical approximation for the effect of input parameter changes of sub-systems can be made. While parameters P which enter multiplicatively in the model just produce an error proportional to ΔP, the situation is a little

bit more complicated for parameters $\mathbb{P}(d)$ like in equation 2.1:

$$\mathbb{P}(d) = ((R-1)\cdot\mathbb{P}(E)+1)\cdot\mathbb{P}(d|\neg E) \quad .$$

There are two cases: if the initial value of $\mathbb{P}(d|\neg E)$ is given, then the errors in $\mathbb{P}(d)$ due to $\mathbb{P}(E)$ and $\mathbb{P}(d|\neg E)$ are multiplicative. But if the initial value for $\mathbb{P}(d)$ is given, then only the error due to this initial value is multiplicative. The error due to a change in the initial value of $\mathbb{P}(E)$ is:

$$\Delta\mathbb{P}(d) = -\frac{(R-1)\cdot\mathbb{P}(E)+1)}{((R-1)\cdot\mathbb{P}(E_{ini})+1)^2}\cdot\mathbb{P}(d_{ini})\cdot\Delta\mathbb{P}(E_{ini}) \tag{4.1}$$

The maximum total error is of course the sum of the absolute magnitude of the individual errors. In a similar manner all equations in the system can be examined for error propagation.

However, this is not sufficient for the analysis of the stability of the system due to the occurrence of feedback loops: the errors have a feedback on themselves and therefore they may accumulate in a geometric series or even worse. Since there is no chance to solve this problem analytically we have to investigate the stability numerically.

As a quantitative analysis two things were done: firstly each input parameter was changed individually. By changing each input consecutively by $\pm 20\%$ the output variables always changed by less than 10%, except when changing the initial pupulation levels. The change was done linearily around the initial value. This is a first indicator that the system is stable, especially in the long time behavior, when the time dependent input parameters do not vary any more (c. f. the standard scenario in the next section).

Secondly the Ljapunov exponent [40] was calculated while some input parameters were varied. The Ljapunov exponent λ measures how fast the trajectories of different initial conditions separate (positive exponent) or converge (negative exponent) in phase-space with time. The definition of the maximal Ljapunov exponent for continuous, differentiable systems is

$$\lambda = \lim_{t\to\infty}\frac{1}{t}\ln\left|\frac{\delta Z(t)}{\delta Z(0)}\right| \quad . \tag{4.2}$$

But since we do not know this movement through phase-space analytically we had to work numerically. The numerical calculation is more difficult and was done in the following way:

A Monte-Carlo sampling of the System Dynamics model was performed: some of the input parameters where varied randomly at the same time. The variations followed a Gaussian distribution with a standard deviation of 20% around their initial values. This was done for 100 different input parameter combinations and the respective results have been calculated. In figure 4.1 we see the results for the variation of all input 'rates', that are those parameters which have the identifier 'rate' in their name. The discrete Ljapunov exponent can be calculated as

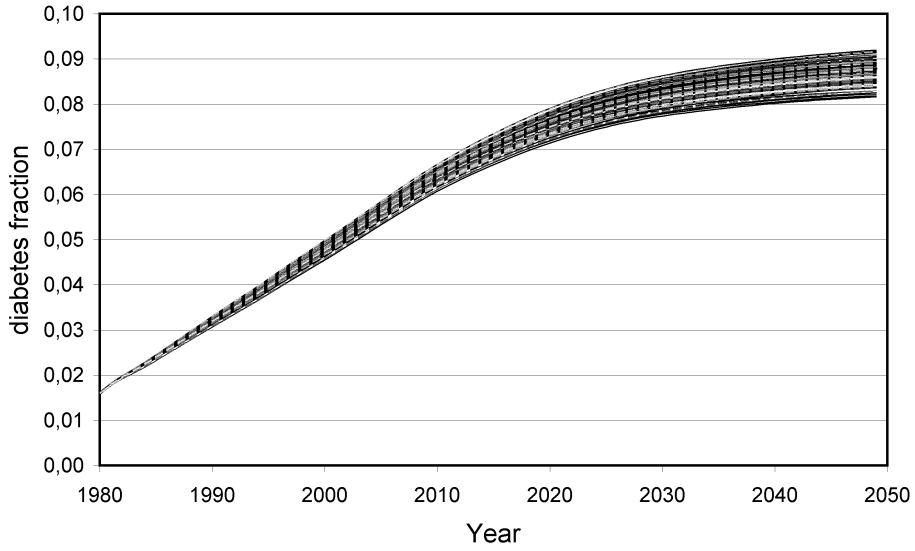

Figure 4.1: The diabetes fraction of the adult population for 100 different input parameter combinations.

$$\lambda = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N} \ln \left| \frac{\partial}{\partial y} f(y(t_n)) \right| \quad . \tag{4.3}$$

Since we do not know the analytical derivate of $f(y(t_n))$ with respect to $y(t_n)$ we need to approximate the differential quotient through

$$\frac{\partial f}{\partial y} = \frac{f(t_{n+1}) - f(t_n)}{f(t_n) - f(t_{n-1})} \quad . \tag{4.4}$$

This approximation is valid since the values for the time step t_{n+1} are calculated from the values of the previous time step. The approximation is quite good as long as the denominator stays away from zero, which is luckily the case. If the simple finite difference quotient $\frac{\Delta f}{\Delta t_n}$ would be calculated instead, a wrong result would be obtained since the effect of the phase-space velocity of the system would be excluded. Now two kinds of averaging have been done:

1. The mean of all runs and the Ljapunov exponent from it have been calculated and

2. the Ljapunov exponents for every run has been calculated and averaged afterward.

For the test runs from figure 4.1 the respective results are $\lambda_1 = -0.0377$ and $\lambda_2 = -0.0373 \pm 0.007$. All of the individual exponents in the second case are negative and the ± 0.007 is the standard deviation. These Monte-Carlo averaged Ljapunov exponents agree very good, the difference is of the order of the standard deviation. For all kinds of variations of the input variables and parameters the Ljapunov exponent turned out to be negative although small (typically between -0.02 and -0.1). There turned out to be no significant difference if 1.000 instead of 100 different parameter combinations were used.
e can now interpret these results: although a variation in the initial conditions leads to a larger difference in the end than in the beginning, as can be seen in figure 4.1, the system turns out to be stable. This can actually be seen in the figure since there is a bend toward a constant limit for all test runs which starts after the year 2010. That there is a bend toward a constant upper limit is a sign of stability.
The conclusion is that the system is quantitatively dynamically stable in accordance with the qualitative expectations, as long as the input data does not leave the compulsory bounds (like probabilities greater than 1 and the like).

4.2 Historical Validation

In this section we give the output data for the standard scenario: all time dependent input variables except the adult population input and the elderly fraction of the population stay constant after the year 2006. First we show that available historical data can be reproduced. And this gives a picture of how the diabetes prevalence in Austria would develop till the year 2050 if the behavior of all influencing factors does not change after 2006. In the next section several possibilities for future development will be discussed.

In figure 4.2 the diagnosed and undiagnosed diabetes and diabetes with complications fractions of the total population are shown. The simulated data agrees to a good degree with the available data. We overestimate the diagnosed fraction compared to the undiagnosed one a little bit because although people can visit medics for free they don't necessarily do it. But the total diabetes cases are in the bandwidth of the available estimations from the references [12], [11], [3] and [41].

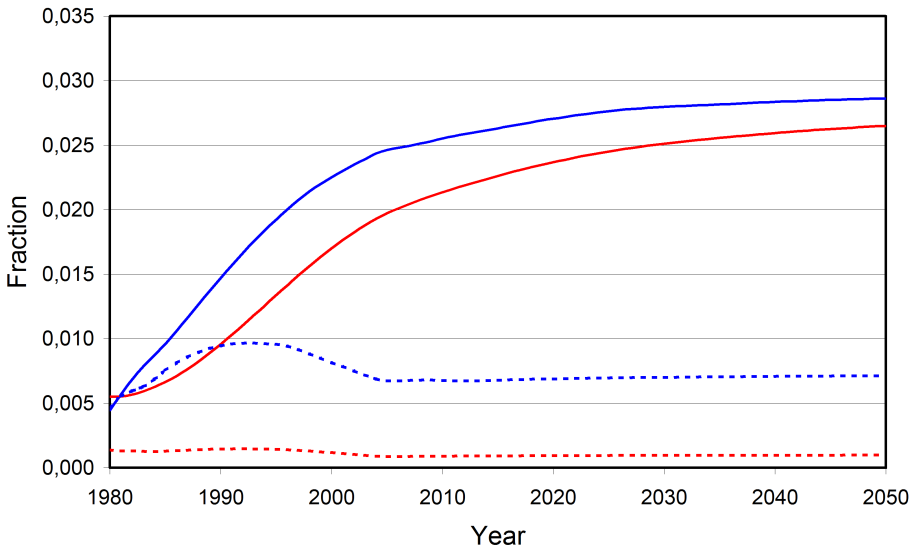

Figure 4.2: The diagnosed (solid) and undiagnosed (slashed) fractions of the total population for the diabetes (blue) and the diabetes with complications (red) populations.

Furthermore we see that the total number of diabetes cases would increase sharply till 2010 while the spread of the disease slows somewhat after that time. This has to to with the age structure of the population. But the absolute number of diabetes cases will continue to grow sharply since the population increases.

In figure 4.3 the prediabetes fractions are shown. We clearly see that the diagnosed fraction only starts to increase after the clinical picture has been introduced in 1994. When we compare this with the data from the US model by J. Homer [23], we see that although diabetes is also on the advance in Austria the initial condition is better and the total rise in not as high as in the USA.

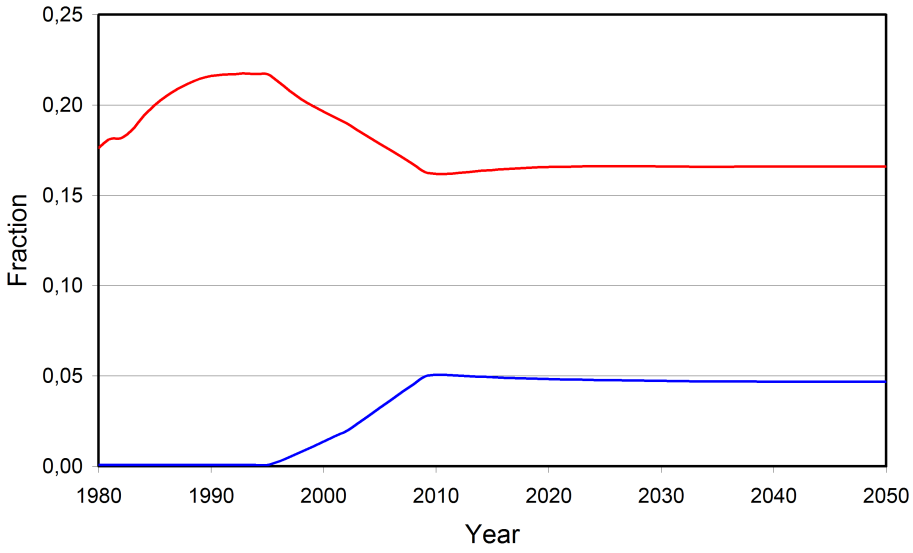

Figure 4.3: The diagnosed (blue) and undiagnosed (red) prediabetes fractions of the total population.

The absolute numbers are omitted since the interest is in the qualitative behavior of the system. Nevertheless we give the total health care costs for prediabetes, diabetes and diabetes with complications in Euro per year in figure 4.4. These refer to the the year 2006 without inflation or other effects influencing prices. We see that the total health care costs increase faster than the continual part of them. This has to do with the much higher costs for complicated diabetes: both, the uncomplicated and the complicated diabetes population, levels increase quite parallel, as can be seen in figure 4.2. But the much higher costs for the complicated level take their toll.

4.3 Experiments

This section describes the set up of four different simulation runs:

1. The step-like change scenario, which is mathematically enlightening,

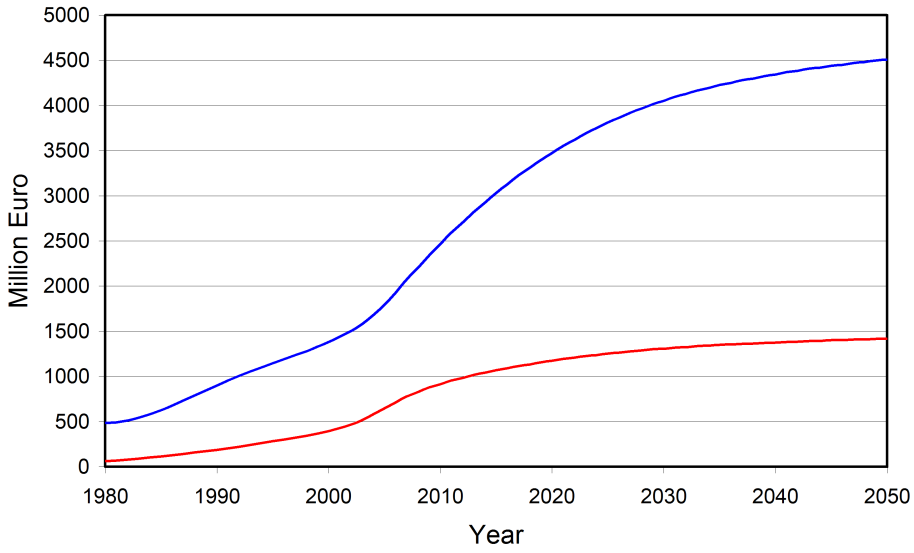

Figure 4.4: The total health care cost (blue) and the continual costs (red) for pre-diabetes, diabetes and diabetes with complications from 1980 till 2050 in Million Euro.

2. the physical activity level (PAL) change scenario,

3. the early detection scenario and

4. the combined scenario.

Note that the simulation runs are always made for the whole population without an distinction with respect to sex. This has two reasons: one is, that the qualitative behavior is always the same for the male and the female test runs since the scenarios are not gender specific. The other reason is that we are only interested in the qualitative behavior.

Step-Like Change

The first simulation run describes how the system responds to a sudden change of the caloric intake. Since the caloric intake in Austria [30] is somewhat below the recommended one [31], we change it suddenly with the year 2007 to the recommended level. The graph of the time series can be seen in the figure 4.5.
From a mathematically point of view it is interesting to see how the system responds to a discontinuous, step-like change in one of the major input parameters. Although the BMI feedback loop is separated from the rest of the system, as can be seen in

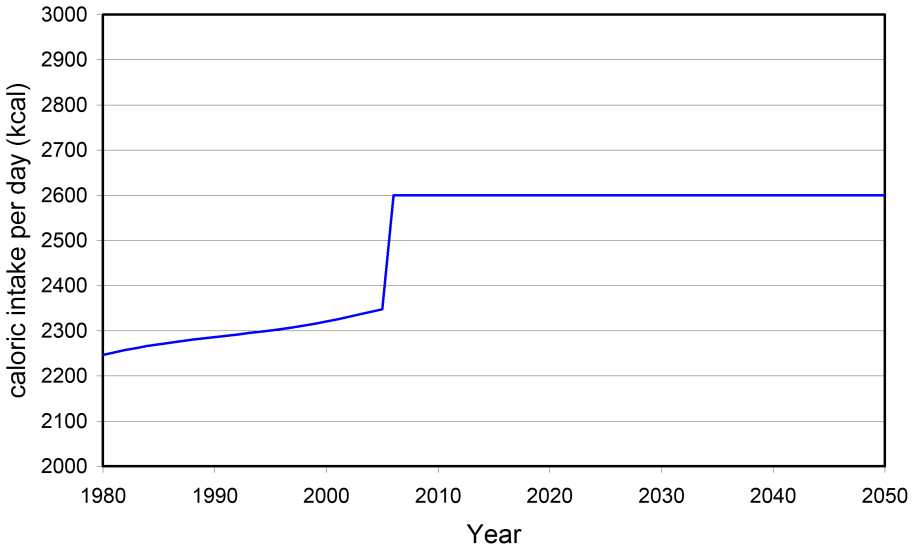

Figure 4.5: The caloric intake from 1980 till 2050 for the step-like change scenario.

figure 2.12 on page 29, the obese fraction of the adult population affects the obese fractions of all levels. So this change is far from trivial.

PAL Change

The physical activity level is in Austria, with values around 1.4, somewhat less than recommended one ([30], [37]). This is partially compensated by the fact, that the daily calories supply is also lower than the recommended value ([30], [31]). In total the BMI, which gives via a lookup function the fraction of obese people, is above the international average.

In this scenario the PAL level is increased over a time of 10 years to the recommended level of 1.65 [31]. This is in accordance with the WHO recommendation of half an hour of physical exercises per day to prevent obesity and type-2 diabetes [2]. This recommendation has been adopted in the Austrian diabetes plan [13]. The 8 years give enough time for educational effects to unfold their impact and the slope is similar to the one in the 1990s. The graph of the PAL over the years is given in figure 4.6.

The question here is how a slow change of the BMI affects the dynamics compared to the sudden increase in the first scenario.

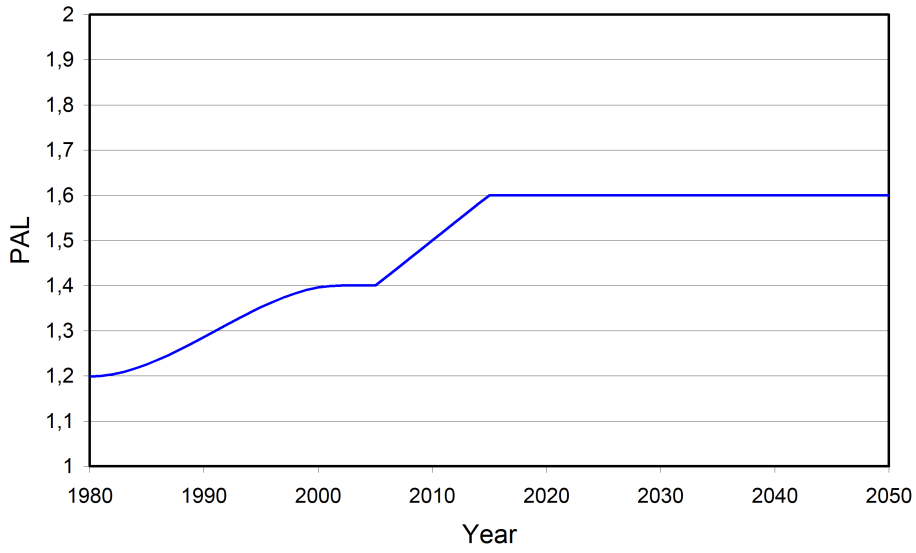

Figure 4.6: The average physical activity level per person from 1980 till 2050.

Early Detection

The early detection measures include the change of time independent parameters. Firstly the average time between two visits to a medical doctor is reduced by one year. Secondly the fraction of high risk people tested for prediabetes is increased, which can be seen in figure 4.7. These measures can be done by an increase in the frequency of the general health screenings and an increased testing of high risk people, for example at the working place. One thing to look at is if more observations of the onset have a measurable effect on the advance of the disease via the disease control.

Physical Activity and Prevention Screening

The last scenario is a combination of the previous two scenarios. The question is whether the two different interventions produce an additional feedback effect.

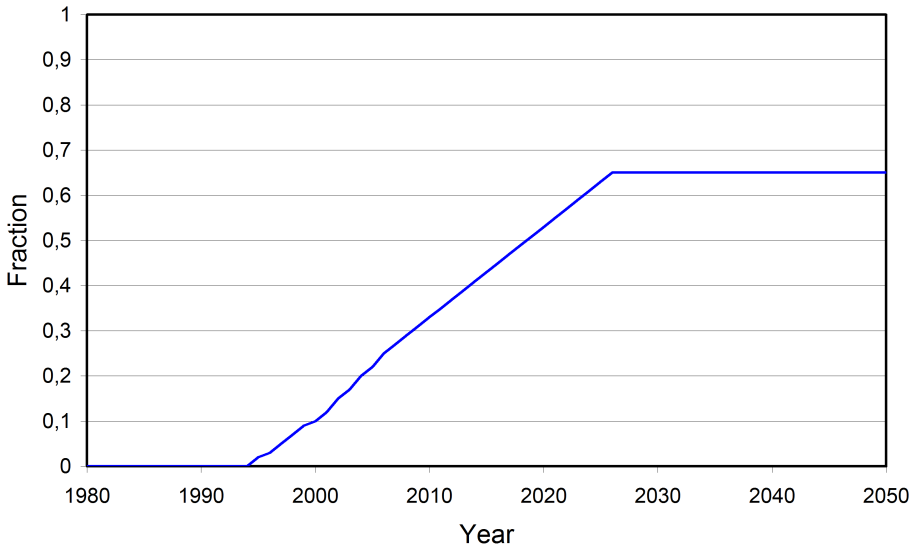

Figure 4.7: The high risks tested for prediabetes fraction of the diabetes testing fraction from 1980 till 2050.

4.4 Results and Interpretation

We have already seen the development for the diabetes prevalence for the standard scenario in section 4.2. In this section we give the results of the other testing scenarios and interpret them.

Step-Like Change

For a sudden increase in the daily consumed calories we get a time delayed increase in the prediabetes and diabetes fraction of the population. This can be seen in figure 4.8. We see that the increase has a delayed onset of about 3 years (increase in 2006, divergence in the diabetes fraction starting 2009). We see that the system responds continuously to a severe discontinuous change of an input variable. In other words the system behaves stable in this context. In figure 4.9 we see the development of the health care costs in this scenario. Again the effect is time delayed. And again we see that the increase in the acute health care costs, which is due to hospital stays, is larger than the increase of the continual costs.

PAL Change

In the second scenario, the the change goes slower and in the other direction. The figures 4.10 and 4.11 again show the total diabetes fraction of the population and the associated health care costs. Although is is of the same magnitude as the step-like change we get a larger reduction than the increase was in the other case. This is due to the feedback mechanism of the PAL concept. Medically speaking physical exercise is much better than making a diet. Another interesting observation is that this policy is especially effective in reducing the uncomplicated diabetes cases.

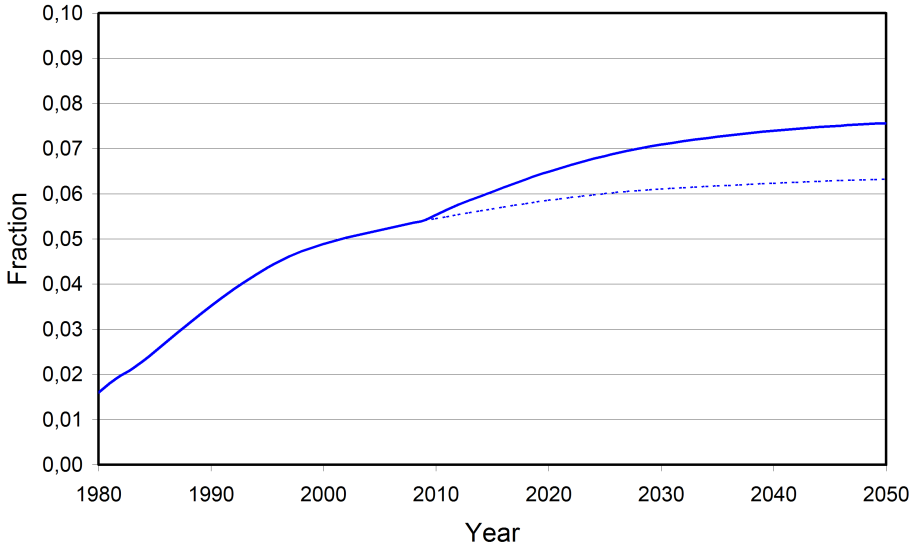

Figure 4.8: The total diabetes fraction of the adult population for the step-like change scenario (solid line) compared to the standard scenario (slashed line).

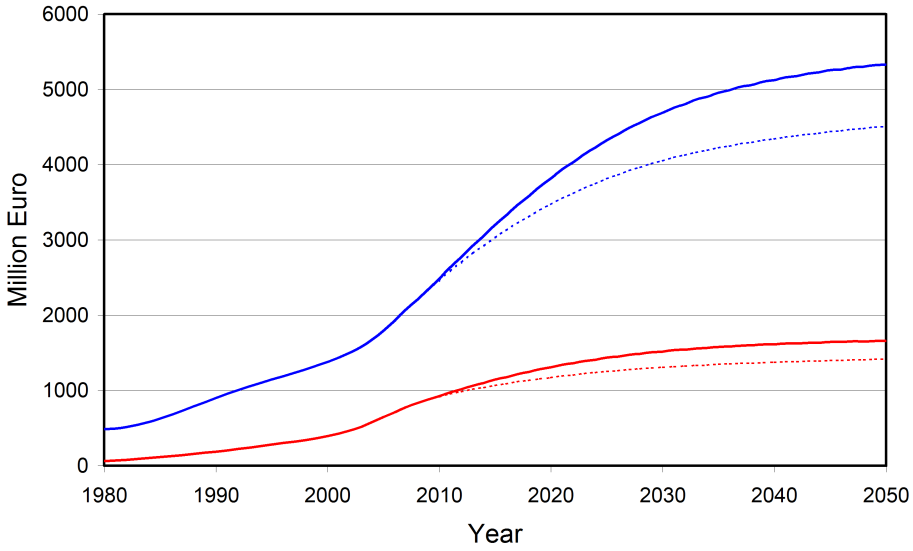

Figure 4.9: The total (blue) and the continual (red) health care costs for the step-like change scenario (solid lines) compared to the standard scenario (slashed lines).

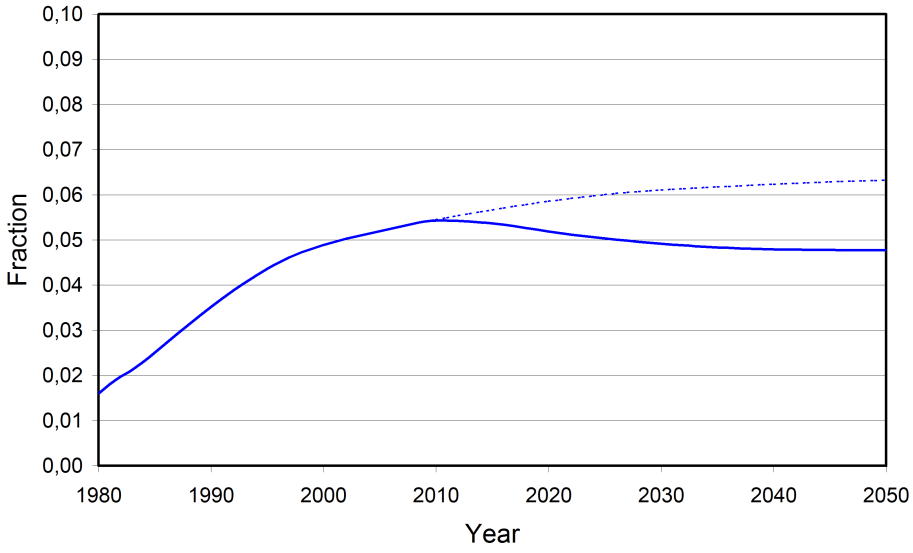

Figure 4.10: The same as figure 4.8 but for the PAL scenario.

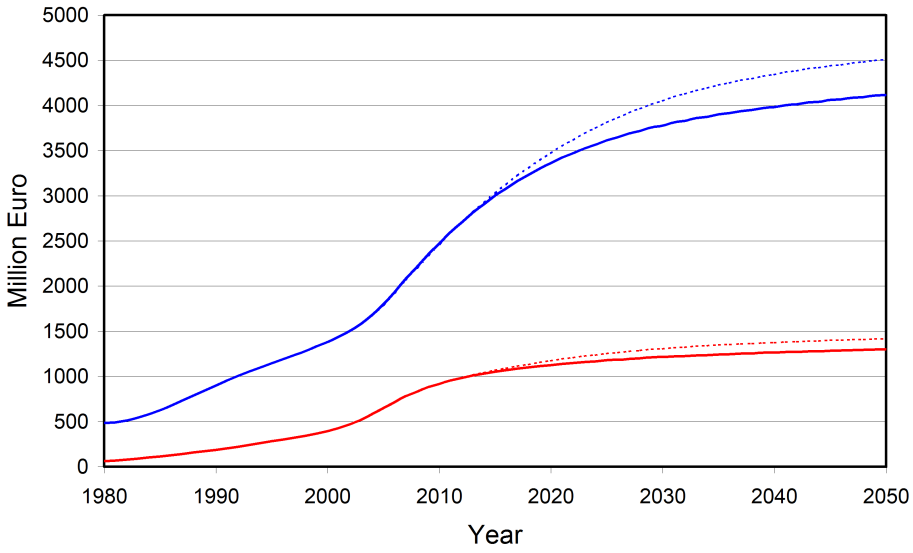

Figure 4.11: The same as figure 4.9 but for the PAL scenario.

Early Detection

If we now increase the number of people tested for prediabetes and shorten the time between medical doctor visits we get the following results: The diabetes fraction of the population decreases slightly, as can be seen in figure 4.12. But the health care costs in figure 4.13 explode ! A rather unsettling and counterintuitive behavior, isn't it ? Now what has happened ? The answer is that the number of diagnosed prediabetes cases increases very much. This leads to a higher fraction of managed prediabetes cases and therefore the costs for the prediabetes management explode. What did we buy with this money ? We bought that the prediabetes recovery rates are almost as high as the onset rates ant that the prediabetes level stays almost constant.

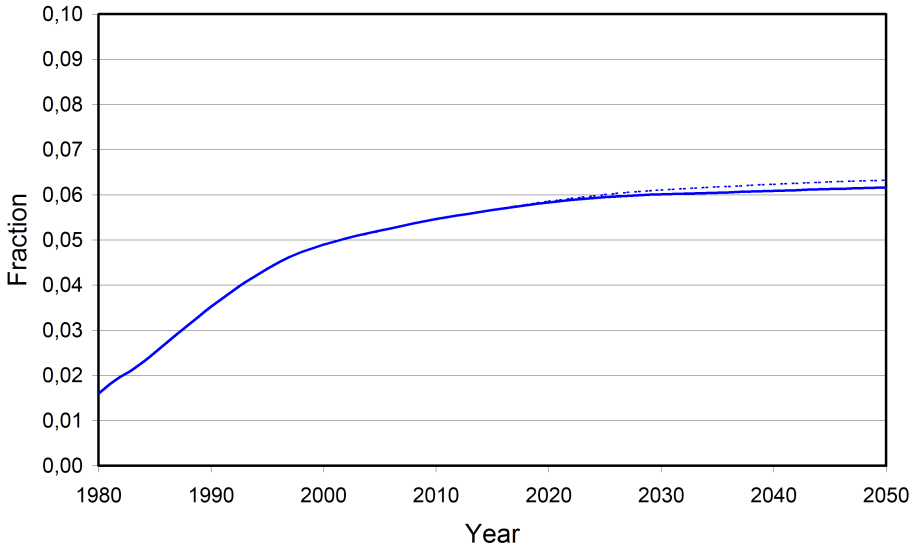

Figure 4.12: The same as figure 4.8 but for the early detection scenario.

Physical Activity and Prevention Screening

If we now combine the measures we get an even better result for the diabetes population, displayed in figure 4.14, while the costs do not increase as much, which can be seen in figure 4.15. The conclusion is that one can start now finding the optimal policy which will lead to less diabetes cases with no additional health care money spent.

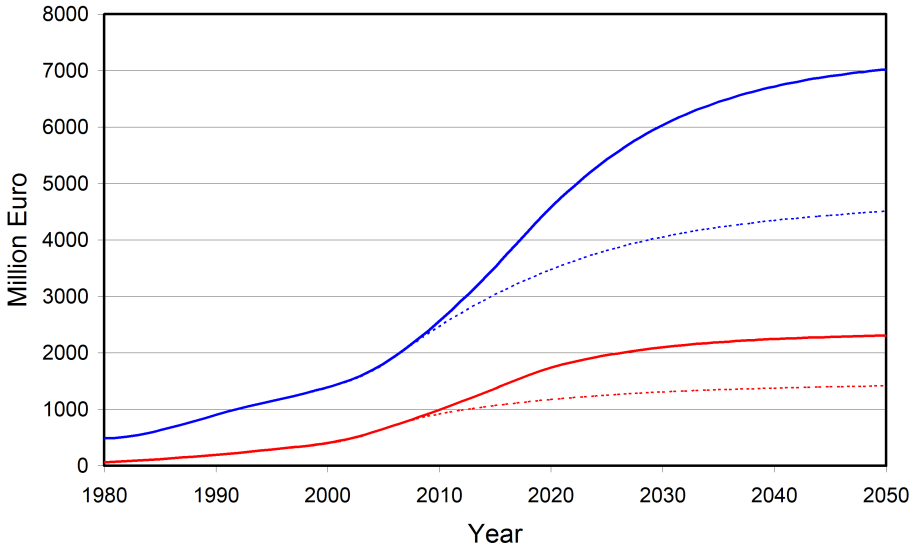

Figure 4.13: The same as figure 4.9 but for the early detection scenario.

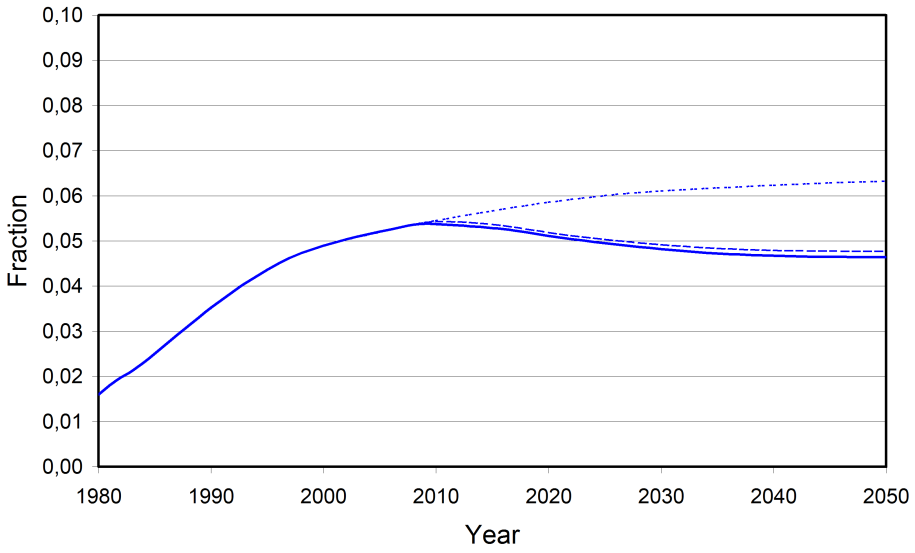

Figure 4.14: The same as figure 4.8 but for the combined scenario. Also displayed is the diabetes fraction of figure 4.10.

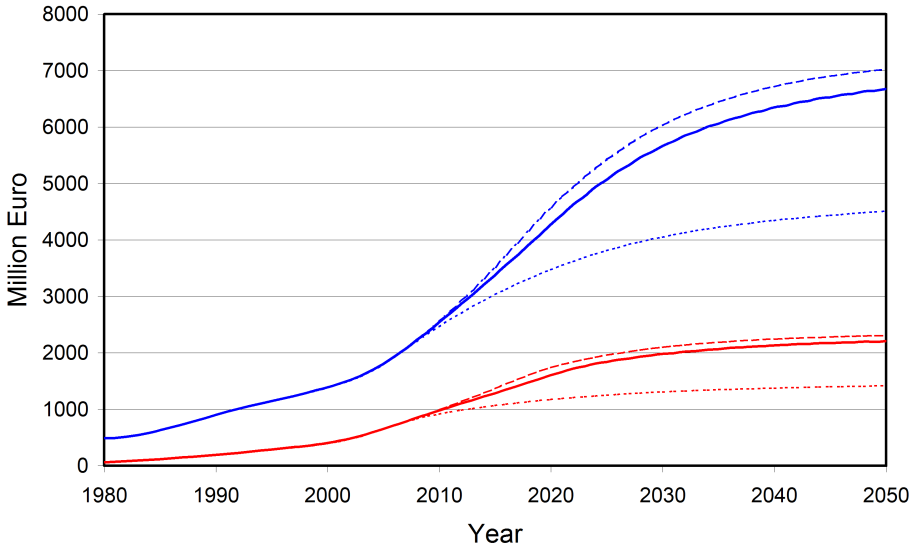

Figure 4.15: The same as figure 4.9 but for the combined scenario. Also displayed are the costs from figure 4.13.

Chapter 5

Conclusion and Outlook

In this thesis we presented a System Dynamics model for the type-2 diabetes mellitus prevalence in Austria. System Dynamics was the method of choice for several reasons:

1. Type-2 diabetes is a progressive chronic disease and levels can be used to describe the different stages of this affliction.

2. Long time scales are involved, since an individual usually stays a long time in the different levels.

3. Diagnosis is always delayed since the onset of the disease occurs unobserved.

4. Unmeasured variables, which are logically necessary like the estimated number of unreported cases, are easily included in the model.

5. There are several risk factors, most prominent age and obesity, affecting the onset and the progression of diabetes.

The iterative modelling of type-2 diabetes leads therefore to a large, very complex system which can not be solved analytically. But System Dynamics is capable to do such simulations numerically.
Also the available input data, originating from different years and of very different quality, suggest a SD treatment: firstly time constant relative risk factors can be obtained from the available studies, secondly some boundary conditions change with time, and finally specialists from different fields of expertise can come and work together on a SD model due to its comprehensible structure.

The model has been developed starting from a model for the USA by J. Homer et al., which is described in reference [22]. Of course the input data is different in Austria, but also structural changes were necessary. The main structural differences of the situation in Austria compared to the USA are:

- the free of charge access to medical treatment in Austria,
- the different structure of the total health care costs and
- a different disease management concept.

They all are due to the different health care systems in Austria and the USA.
Furthermore we included the concept of the physical activity level instead of the calories spent by physical exercises per day. This creates additional feedback loops. And a distinction by sex was made since the risk factors are very different and other policies are necessary. This was possible only because evaluable data exist for Austria. The collection of the required input data has been one of the major tasks solved for this work.
In this thesis it was shown that the SD model for type-2 diabetes mellitus is numerically stable, a previously unpublished result. It is capable to reproduce the right historical data which have been available. Furthermore the following conclusions are drawn from the different simulation runs in chapter 4:

1. Type-2 diabetes will rise significantly over the next 20 years. There is virtually no possibility to achieve a short-term relief due to the long time delays involved.

2. Half an hour of physical exercise per day is a good recommendation of the WHO to prevent type-2 diabetes since it reduces the prediabetes prevalence by 20% and the total diabetes prevalence by 30% as seen in figure **??**.

3. Early diagnoses and good control can ease the burden of diabetes and prevent harmful consequences. It can reduce the health care costs and therefore it can play an important role in the sustainable financing of the Austrian social insurance system.

4. Both of the last two measures together would have an increased positive effect on the diabetes prevalence and the total health care costs.

Recapitulatory we can say that the SD diabetes model is a useful tool for the evaluation of different health care policies for large populations, where field studies simply are impossible.
Future work can be done into two different directions: one may further improve the structure of the model or one may apply the model to other populations.
Structural considerations include:

- combination of the diabetes model with a model for obesity since there are very strong correlations,
- simulation of the type-2 diabetes prevalence in combination with other methods in a hybrid modelling approach and
- development of a testing environment to analyze the stability of this (and other) models.

The application of the model to other populations may go in the following directions:

- The comparison of different provinces in Austria with the idea to compare the effectiveness of different health care policies and by this reduce the costs of case studies and get faster results. We should always have in mind that we are focusing on long term effects with a mean reaction time of up to 20 years.
- Application to people with a different social background to find tailored intervention measures which have the greatest impact.

Type-2 diabetes mellitus constitutes a major public health care challenge but I hope this work helps to find the best way to manage this problem.

Acknowledgment

I'd like to thank my family who supported me in going for a second master degree since they recognized earlier than I did that mathematics is my true calling.
Then I say tanks to my advisor, Felix Breitenecker, who reminded me that science can be fun and that autonomous work leads to greater insights.
I'm also very thankful for the guidance and support of Günther Zauner and Niki Popper, who gave me an introduction to modelling, and to Florian Judex, who gave me the introduction to the modelling and simulation community.
And finally I thank all my friends for their moral support and many hours of fun, especially I'd like to say to the 'Fachschaft':

Go (party) on !

Bibliography

[1] A. Jones et al. , 'Understanding Diabetes Population Dynamics Through Simulation Modeling and Experimentation', American Journal of Public Health, 96(3), 488-494, 2006

[2] World Health Organization, Diabetes fact sheet, URL: http://www.who.int/mediacentre/factsheets/fs312/en/index.html , September 2006

[3] Federal Ministry of Health and Women, EU Conference on Prevention of Type 2 Diabetes, Conference Report, Vienna, April 2006

[4] Joseph L. McCauley and Cornelia M. Kffner, 'Economic system dynamics', Discrete Dynamics in Nature and Society, vol. 2004, no. 1, 213-220, 2004

[5] D. Evenden, P. R. Harper, S. C. Brailsford and V. Harindra, 'SystemDynamics modelling of *Calmydia* infection for screening intervention planning and cost-benefit estimation', IMA Journal of Management Mathematics **16**, 265-279, 2005

[6] Wikipedia, URL: http://en.wikipedia.org/wiki/Diabetes_mellitus , April 2007

[7] World Health Organization, ICD-10, diabetes mellitus, URL: http://www.who.int/classifications/apps/icd/icd10online/?ge10.htm+e10 , 2007

[8] World Health Organization, 'Diabetes and Genetics', URL: http://www.who.int/genomics/about/commondiseases/en/index.html , May 2007

[9] Initiative Arznei & Vernunft, 'Typ 2 Diabetes', Wien, 08/2004

[10] 'Definition, Diagnosis and Classification of Diabetes Mellitus and its Consequences, Part 1: Diagnosis and Classification of Diabetes Mellitus', World Health Organization, Geneve, 1999

[11] A. Rieder et al. , 'Österreichischer Diabetesbericht 2004', Altern mit Zukunft, Vienna, 2004

[12] 'Screening for Type 2 Diabetes', World Health Organization, Geneve, 2003

[13] Federal Ministry of Health and Women, 'Diabetes mellitus - a challenge for health policy', Vienna, February 2006

[14] R. Wagner, 'Vermittlung Systemwissenschaftlicher Grundkonzepte', Diplomarbeit, Karl-Franzens-Universität Graz, Berlin, 2002

[15] Frank Schöneborn, 'Strategisches Controlling mit System Dynamics', Physica Verlag Heidelberg, 2004

[16] D. Meadows, J. Randers, D. Meadows, 'Grenzen des Wachstums - das 30-Jahre-Update', Hirzel Verlag, 2006

[17] H. S. Bierman, L. Fernandez, 'Game Theory with Economic Applications', 2nd edition, Addison-Wesley, 1998

[18] Amanda A. Honeycutt et al. , 'A Dynamic Markov Model for Forecasting Diabetes Prevalence in the United States through 2050', Health Care Management Science 6, 155164, 2003

[19] C. Beauchemin, S. Forrest and F. Koster, 'Modeling influenza viral dynamics in tissue.', The 5th International Conference on Artificial Immune Systems (ICARIS), Volume 4163/2006, 2006

[20] Catherine Beauchemin, 'Probing the Effects of the Well-mixed Assumption on Viral Infection Dynamics', J. Theor. Biol., 2006

[21] Paul Watzlawick, 'Selbsterfüllende Prophezeiungen' in 'Die erfundene Wirklichkeit' (edited by Paul Watzlawick), Serie Piper, Munich, edition June 2006

[22] J. Homer, A. Jones, D. Seville, 'Diabetes System Model Reference Guide', Sustainability Institute, USA, October 2004

[23] J. Homer et al., 'The CDC's Diabetes Systems Modeling Project: Developing a New Tool for Chronic Disease Prevention and Control', 22^{nd} International Conference of the System Dynamics Society, Oxford, England, July 2004

[24] J. Homer et al. , 'Obesity Population Dynamics: Exploring Historical Growth and Plausible Futures in the U.S.', submitted to the 24^{th} International System Dynamics Conference, Nijmegen, July 2006

[25] Abdel-Hamid, 'Modeling the dynamics of human energy regulation and it's implications for obesity treatment', System Dynamics Review 18(4), 431-471, 2002

[26] 'Steirischer Seuchenplan', Das Land Steiermark, O. Feenstra (Hrsg.), Graz 2002, Update März 2005

[27] Statistisches Jahrbuch Österreichs 2006, STATISTIK AUSTRIA, 2006

[28] Datenbank ISIS, URL: http://www.statistik.at/isis/current/isis_gui.shtml , STATISTIK AUSTRIA, 2007

[29] I. Elmadfa, P. Burger et al. , 'Österreichischer Ernährungsbericht 1998', Wien, 1998

[30] I. Elmadfa, H. Freisling, J. König et al. , 'Österreichischer Ernährungsbericht 2003', 1. Auflage, Wien, 2003

[31] 'Referenzwerte für die Nährstoffzufuhr', Deutsche Gesellschaft für Ernährung (DGE), 1. Auflage, 2. korrigierter Nachdruck, Frankfurt am Main: Umschau/Braus, 2001

[32] NAHES Datenbank ,Hauptverband der Sozialversicherungsträger, Wien, 2007

[33] Demographisches Jahrbuch 2005, STATISTIK AUSTRIA, Wien, 2006

[34] E. Pochobradsky et al. , 'Soziale Ungleichheit und Gesundheit', Sektion IV/A/5, Vienna, 2002, Nachdruck 2004

[35] 'Therapie Aktiv', Disease Management Program, Steiermärkische Gebietskrankenkasse,2006

[36] E. Bachinger, M. Csitkovics, K. Waais et al. , 'Gesundheitsbericht Wien 2004', Stadt Wien (Hrsg.), Wien, 2005

[37] I. Elmadfa, J. Blachfelner, H. Freisling et al. , '2. Wiener Ernährungsbericht 2004', Stadt Wien (hrsg.), Wien, 2005

[38] Mikrozensus, 'Gesundheitszustand der Österreicher', STATISTIK AUSTRIA, Wien, Dezember 1991

[39] Statistische Nachrichten 6/1994, STATISTIK AUSTRIA, Wien, 1994

[40] A. Katok, B. Hasselblatt, 'Introduction to the Modern Theory of Dynamical Systems', Cambridge University Press, 1995

[41] H. King et al. , 'Global Burden of Diabetes', Diabetes Care 21, p. 1414-1431, 1998

[42] ORF ON Science, 'China: Gesundheitskrise durch Diabetes?', URL: http://science.orf.at/science/news/146700 , 14. 02. 2007

[43] E. Urbas, J. Klimont, E. Bachinger et al. , 'Chronische Krankheiten in Wien', Stadt Wien (Hrsg.), Wien, 2004

[44] M. Klampfl, 'Diabetes mellitus Typ-2 Screening: Eine Übersicht über Methoden, Screening Tests und Empfehlungen', Graz, 2000

[45] 'Lebenserwartung und Mortalität in Wien und Österreich Internationaler Vergleich', Stadt Wien (Hrsg.), Wien, Mai 2003

[46] W. Freidl, W. J. Stronegger, Ch. Neuhold, 'Lebensstile in Wien', Stadt Wien (Hrsg.), Wien, 2003