1 Introduction

1.1 General remarks about modelling

1.1.1 An introductory example

As is well known, each year in winter, influenza A strikes across almost all countries and regions of the world. It usually begins with a few cases and can develop into a real epidemic. After some time there are again only a few people who are still ill and finally the infection dies out completely. We also know that some individuals do not get infected at all. At the beginning of such an epidemic we might be interested to know how many people will get infected and how long it will take until there is no case left. Medical officers from the national health services or researchers from universities might make first estimates based on experience of previous epidemics. But such predictions can be very unprecise, if one does not think about what is really happening during such periods.

This is where mathematics can help with models: mathematical models make it possible to describe very accurately the behaviour of systems which are too complicated to describe in natural language and to predict the development in the future. The language of mathematics (differential equations for example) is universal in the sense that we can describe processes of physics, chemistry, biology, medicine, engineering and economics. But these models have to be built first; so after a phenomenon of the real world has been described in natural language, we have to translate the information into mathematical language. This translation is called modelling. As soon as the model has been set up, the whole apparatus of mathematics can be used to analyse the mathematical model. The conclusions we derive must later be retranslated into natural language.

Let us examine the modelling process using the example from above: the development of the influenza-epidemic in time. It begins with medical officers having to explain to a mathematician in detail what is happening during the time that such an epidemic spreads. This is a very important stage in the process. As Barbour (1989) writes (in a more general context):

> "One great advantage of including a mathematician as collaborator in a project team is his ignorance. He is not there to know, but to think. It will be necessary to explain to him in precise detail what the problems are, and what sorts of solutions are being considered, if any. In the course of the explanation, it will prove necessary to acquaint him with a lot of knowledge which the experts in the relevant fields of research take for granted. By the time all his questions have been answered, and he feels that he understands the problem fully, whoever has had to do the explaining will also understand the problem very much better."

In our example, the problem could be described in the following way: A very small part of the population (of the country of interest) is infected with a virus and can potentially infect other people. These people might have got infected through travelling into distant countries or through contact with people coming from such countries (tourists, businessmen and businesswomen, refugees). They can infect other people if the contact between two people is close enough and the potentially newly infected person is susceptible to the virus. Then the newly infected people can infect other people too. But after some time people recover and can no longer infect other people and themselves are no longer susceptible.

This might be the explanation of the medical officers. The facts stated above are well accepted in medicine. However, in other more difficult research problems, it might be that such "facts" should not really be taken for granted. Again Barbour (1989) writes:

> "There is another aspect of the mathematician's lack of knowledge which can at times also prove fruitful, and that is his consequent lack of prejudice. To the expert, certain ideas are accepted as true, on the basis of received wisdom, and others as impossible. If this were not the case, it would be impossible to learn from other people's experience, and science as we know it would be useless. However, not all scientific theories are accepted as a result of overwhelming supporting evidence. Quite often, a theory is tentatively advanced on the basis that it explains some experimental results reasonably well and fits in with other comparable theories, and then gradually becomes accepted if no-one publishes conflicting evidence, even when no strongly corroborative evidence comes to light. The effect of being referenced frequently in subsequent papers of itself lends considerable weight to a theory, even when the reference is not intended as direct support: it serves as evidence of acceptance. Thus an expert tends to possess a large amount of knowledge of varying soundness, and has a difficult task distinguishing between the true and the not so true. The mathematician, on the other hand, knows none of the theories, and bases his judgements primarily on the data to hand, taking the theories which have been explained to him as clearly defined assumptions, rather than as facts. This can be of substantial benefit when the data are found to be in conflict with theory. This is not, of course, to suggest that the mathematician can dispense with the expert: far from it. Rather, it shows how, in a perhaps unexpected way, the mathematician can be of use to the expert."

After having received an explanation in natural language of the process that leads to the spread of the virus, the mathematician wants to model this process. We make the following assumptions, translating bit by bit the above explanation into mathematical language.

We introduce the three variables $S(t), I(t)$ and $R(t)$ which stand for the proportion of the entire population of susceptibles $(S(t))$, infectives $(I(t))$ and removed individuals $(R(t))$ at time t. Some remarks should be made. By susceptibles we mean individuals who can get infected with the virus if they are unlucky; by infectives we mean people who are ill and can still infect

other people; the removed individuals are people who have either been immune from the beginning, or people who have acquired immunity through having recovered from the infection, or people who have been isolated or died. So obviously we have at every moment in time that $S(t) + I(t) + R(t) = 1$ and $0 \leq S(t), I(t), R(t) \leq 1$. We now introduce two parameters: α and β, $\alpha, \beta > 0$. They denote the rate at which people make potentially infectious contacts (α) and the rate at which people recover (β). An infection can only occur if an infected person has a close enough contact with a susceptible person, so the rate at which the proportion of susceptibles decreases at time t must be $\alpha S(t)I(t)$. On the other hand only infected people can recover from the infection, so the increase of the proportion of removed people at time t takes place at a rate of $\beta I(t)$. These statements can be written out in a system of differential equations:

$$\begin{aligned} \frac{dS(t)}{dt} &= -\alpha S(t)I(t) \\ \frac{dI(t)}{dt} &= \alpha S(t)I(t) - \beta I(t) \\ \frac{dR(t)}{dt} &= \beta I(t). \end{aligned} \tag{1.1}$$

The derivatives on the right side sum up to 0 which should be as $S(t) + I(t) + R(t) = 1$ for all t. A reader not used to modelling concepts might not fully understand the meaning of the expression "rate", which was used when α and β were introduced. It is not easy to explain the precise meaning of these rates (but in 1.1.2 the rates can be precisely defined in a stochastic environment). But medical researchers should be aware of the fact that by using such rates we automatically made two important assumptions. First, we assume that at any time the infectives and susceptibles mix infinitely fast so that a particular person has uniformly the same chance to meet any other person in the population. The second assumption makes clear how we suppose that people pass on from being infective to being removed. In our model, we assume that looking at the proportion of infectives at a time t_0, there is a half life so that after such a half life only half of those infective at time t_0 are still infectives (although new ones might have turned up) and then again a half life later only one quarter of them and so on. This assumption might be questioned because researchers might suggest that for maybe three or four days almost everybody stays infective if he has just become infected.

After having built this model, we are going to analyse it. Looking at the second equation in (1.1) and assuming that $I(t) > 0$ we see that the number of infectives increases if $\alpha S(t) - \beta > 0$ and decreases if not. Suppose that $S(t)$ is almost 1 at the beginning of the potential epidemic outbreak (so $I(t)$ and $R(t)$ must be close to 0, meaning that there are only very few people introducing the virus to the population and almost nobody is immune). Then

the number of infectives increases if $\alpha - \beta > 0$ or $\alpha/\beta > 1$ and decreases if $\alpha/\beta < 1$. It seems, that α/β is critical to the development of this epidemic. Let us denote this ratio by R_0; so $R_0 := \alpha/\beta$. It can be shown that R_0 is the average number of people an infected person infects himself until he recovers (under ideal conditions for the spread of the epidemic, that is if every contact of an infected is with a susceptible (i.e. if $S(t) \doteq 1$)). In the general theory of mathematical epidemics, this ratio is called the basic reproduction ratio. In this model we see that one answer to the questions the medical officers pose is that if $R_0 < 1$ the epidemic does not develop and only a few more people get infected. If $R_0 > 1$ the epidemic spreads and the peak of the infection is reached as soon as $\alpha S(t) - \beta = 0$. It is clear that the only stationary solutions are the trivial solutions $(S, I, R) = (c, 0, 1 - c)$ where $c \in (0, 1)$. The development in time can be easily simulated using a computer. Two simulations with different values for $(S(0), I(0), R(0))$ have been carried out (using "Mathematica") and are presented in Diagrams 1.1 and 1.2.

Diagrams 1.1 and 1.2: Development of $S(t)$, $I(t)$ and $R(t)$ with different initial values (see text below) in time.

The first diagram shows a typical situation where $S(0) = 0.9, I(0) = 0.01$ and $R(0) = 0.09$ and the parameters have the following values $\alpha = 5, \beta = 3$. We have a full outbreak and less than 40% ($S(\infty)$) of the susceptibles do not get infected with influenza A at all. The second plot shows the situation where a vaccination program has been carried out and $S(0) = 0.66, I(0) = 0.01$ and

$R(0) = 0.33$, with α and β as above. One third of the population has been vaccinated. There is still an outbreak; but comparing the curve for $I(t)$ in both situations the vaccination program has helped quite a lot. If more than 40% of the population had been vaccinated the epidemic would have levelled off very fast because then $\alpha S(0) - \beta < 0$ from the very beginning and so the proportion of infectives would have gone to zero very fast. So these simulations give us an idea of how such an infection develops.

This is only an introductory example and we do not want to spend too much time on it. But we want to show how one could proceed. We might want to check if this model is close enough to reality. So we need data. We have to know how many people are immune from the very beginning ($R(0)$). This is different from case to case depending on the particular virus and the genetical and immunological situation of the population examined. The removal rate (β) too is dependent on the factors above and the medical care which ill people receive. The contact rate (α) depends on how easily the infection can be carried on from one person to another and on how mobile people are. One might try to estimate such parameters or take values measured from recent influenza outbreaks or from other countries.

If the predicted development is similar to the actual development, countermeasures could be discussed using model (1.1): β is not easily influenced except if infected people are ordered to stay in bed or isolated in another way. Before an infection develops people could be vaccinated. This increases $R(0)$ and decreases $S(0)$. As the increase of the proportion of infectives is governed by $\alpha S(t) - \beta$ this could prevent an epidemic developing. Individually people could try to prevent getting infected by not going out much, avoiding public transportation, or other measures; this would decrease α if the measures are successful.

If the predicted development of a mathematical model does not at all mirror the behaviour in the real world, the assumptions must be changed. But mathematical models must not be made too complicated because otherwise the mathematics becomes too complicated and/or the results can no longer be interpreted. The moral is: **MAKE MATHEMATICAL MODELS AS SIMPLE AS POSSIBLE!**

A medical researcher might ask how high the probability is that the epidemic develops if a single person introduces the virus to the population. We know that in reality it is quite possible, that even if the environment is perfect for the spread of an epidemic, simply by chance the infected person might stay at home or not have any infectious contacts for other reasons. For example, the infected person might go to the mountains with a group of people who all get infected, but when they return all have recovered. Questions like these cannot be answered using model (1.1). For this we need a new model using a completely different approach.

1.1.2 Stochastic and deterministic approaches

To further motivate this section let us quote Barbour (1989):

> If the world behaved in exact accordance with simple deterministic mathematical equations, as is apparently the case in certain laboratory experiments in physics, experimental data would give immediate and easily interpretable information about the underlying system. In general (...) there may be insufficient information available about certain features of the system for an exact model to be usable. In such cases, it frequently proves fruitful to describe those parts of the system whose behaviour cannot be accurately specified by allowing for the possibility of 'random' behaviour. The first such models were used to describe the frequencies of various outcomes of games of chances. Of course, the throw of a die is in principle a perfectly well determined physical system, in which, with the right information about geometry, momentum, position, elasticity and the like, the outcome of a given throw can be exactly predicted. In practice, the necessary information is simply not available, and exact prediction is not possible: in contrast, however, the fictional mathematical description of the outcome as the result of a simple random process turns out to be extremely useful. Almost every observation concerned with the real world contains an element which it is convenient to describe as random.

We now build a model which allows for randomness, that is, a stochastic model. The drawback of the stochastic models is that it is no longer possible to present the development over time in such a compact form as we could by showing the simulated results (Diagram 1.1) for the deterministic model.

Let $i(t)$ denote the number of infective people at time t and assume that the number of susceptibles is so large that we can take it to be ∞. We are only interested in whether the number of infected people develops at the beginning or if the infection dies out immediately. We assume that the population is mixing infinitely fast and any contact an infective makes is therefore with a susceptible person. Suppose all infective people make infectious contacts according to a Poisson process with rate λ (if we want to include in the model that a proportion c, where $c \in (0,1)$, of the population is immune from the very beginning, the rate of the Poisson process can be altered to be $(1-c)\lambda$). Once a person has been infected he or she stays infective for a length of time which is exponentially distributed with parameter μ. For an introduction to Poisson processes and exponential distributions see Ross (1983). We have then the following transition rates in model (1.2):

$$\begin{aligned} i(t) &\to i(t)+1 \text{ at rate } \lambda i(t) \\ i(t) &\to i(t)-1 \text{ at rate } \mu i(t). \end{aligned} \tag{1.2}$$

At time t events occur at a total rate of $i(t)(\lambda+\mu)$; the probability that such an event is a new infection is $\lambda/(\lambda+\mu)$; the probability that a person recovers and is not infective anymore is $\mu/(\lambda+\mu)$. Intuitively one might suggest here

that R_0 could be defined to be λ/μ. In fact this is the ratio that separates growth from extinction. R_0 is again the average number of people an infected person infects himself until he recovers (under ideal conditions for the spread of the epidemic, that is if every contact of an infected is with a susceptible). But in this stochastic model it is not correct to say that if $R_0 > 1$ then the epidemic develops. It may well be that the infection dies out right at the beginning by chance even though R_0 is much greater than 1. The motivation for building such a model came exactly from considering the question "what is the probability that the epidemic dies out right at the beginning even though the environment is suitable for spread?" We note that this probability can be calculated. Given a single person is infected right at the beginning, the probability that the infection dies out may be $q \leq 1$. Since we assumed that people act independently, the probability that the infection dies out starting with n infected people is q^n. So q can be calculated as follows. With probability $\mu/(\lambda + \mu)$ the first event to occur is that the infective person recovers. Then the danger is over. But, on the other hand, with probability $\lambda/(\lambda + \mu)$ an infection occurs. Then there are two infective people. But we know, that with two infective people the probability that the infection dies out is q^2. So we have the equation

$$q = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} q^2.$$

We get $q = 1$ if $\mu \geq \lambda$ and $q = \mu/\lambda$ if $\lambda \geq \mu$. As can easily be seen, q gets smaller if λ gets larger relative to μ as expected. So here we have the answer to the question above!

The stochastic and the deterministic approaches are the most common ways to model real life phenomena. In chapter 3 of this thesis we show how stochastic models and corresponding deterministic models are connected to each other. The main results are a law of large numbers where the stochastic model behaves more or less like the deterministic model if the number of objects studied is large and the fact that the expectation of the stochastic process studied is near to (or even exactly) the solution to the corresponding differential equation.

Remark: It must be made clear that the expectation of $i(t)$ is certainly not $I(t)$. The first (deterministic) model (using $I(t)$) is not the corresponding deterministic model of the stochastic model (using $i(t)$). The stochastic model was introduced to model the initial phase of the epidemic. In fact we have $\mathbb{E}[i(t)] = i(0)e^{(\lambda-\mu)t}$ and $i(0)e^{(\lambda-\mu)t}$ is the solution to the deterministic model defined by the differential equation

$$\frac{di}{dt} = \lambda i - \mu i$$

which is certainly not system (1.1).

This simple introductory example and the general remarks about modelling suggest the way that we proceed in this thesis. In 1.2 we explain *in natural language* the mechanisms that govern the spread of the disease schistosomiasis. In 1.4 we *translate these mechanisms into mathematical language*, that is build the models. In chapter 2 the stochastic models are analysed. In chapter 3 the stochastic and the deterministic approaches are linked in various ways and in chapter 4 the deterministic models are analysed. Additionally in chapter 5 we present some results from simulations.

If the reader is interested in the history of mathematical models used in epidemiology, an introduction can be found in Bailey (1957).

1.2 Medical and biological information about schistosomiasis

1.2.1 An overview

Schistosomiasis (bilharzia) is a severe tropical disease. It is endemic in most regions where it occurs, that is in parts of Asia, Africa and South America. Estimates of how many people are infected vary from 200 to 400 million (1997). It was described by the ancient Egyptians 3,000 years ago (they mentioned blood in the urine as a symptom) and Napoleon's troops suffered terribly from this disease during the Egyptian campaign. The worm (schistosome) which causes this disease was first discovered by the German doctor Theodor Bilharz (1825 - 1862), hence the German name "Bilharziose" for the disease. The whole cycle was first completely understood at the beginning of this century.

1.2.2 Description of the cycle

Schistosomiasis is caused by schistosomes. A schistosome is a special type of worm that belongs to the class of trematodes. There are various types such as Schistosomiasis haematobium (Africa, Arabic countries and India), S. mansoni (Africa, Arabian peninsula, Caribbean and South America), S. japonicum (Southeast Asia and East Asia), S. intercalatum (West Africa). The whole infection process includes a definitive host and an intermediate host. The definitive host is usually a human being in whom sexual reproduction takes place. But depending on the type of schistosomiasis it may be any mammal (S. japonicum). The intermediate hosts are snails of the types Bulinus, Biomphalaria and Oncomelania (depending on the type of schistosomiasis) who live in fresh water. The reproduction in the snails is asexual. The whole process consists of the following parts (as understood in 1999):

1 Worm in blood vessel of man (×4)

2 Egg (×130)

3 Miracidium (×145)

4 Intermediate host (snails)

5 Larva stages in snails (×16)

6 Cercaria (×45)

Diagram 1.3: Cycle of transmission of bilharzia (here schistosomiasis mansoni).

Male and female schistosomes mate in the blood vessels of the large intestine or bladder of man and produce eggs. Schistosomes can live for several years (estimates range from 2 to 20 years). The eggs are deposited in the bladder (90 %) or the large intestine (10 %). They are released with urine or faeces. If the urine or faeces are released into a well-disposed environment, the eggs can develop. This "well-disposed environment" means that the eggs are required to fall into fresh water, if possible standing or otherwise with not too high a current, at a temperature of at least 20^o C. Mineral conditions should not be too extreme and the snails mentioned above must live in the same water. Then the eggs hatch and become miracidia. These miracidia must penetrate the snails in one or two days. Through diverse stages of asexual reproduction inside the snails finally after approximately 5 weeks so-called cercariae leave the snails and swim around in the water. If they find a man in the next 1 or 2 days they penetrate his skin in only a few minutes. After some 4 to 12 weeks (depending on the type of schistosomiasis) the now adult worms reach the blood vessels of the large intestine and the bladder. Then they mate again completing the cycle.

In schistosome mating the thicker male embraces the female (see Diagram 1.3). The whole body of the male looks "split". This is where the disease got its name from: Schistosome is $\sigma\chi i\zeta\epsilon\iota\nu\sigma\tilde{\omega}\mu\alpha$ in Greek, where "$\sigma\chi i\zeta\epsilon\iota\nu$" means "split" and "$\sigma\tilde{\omega}\mu\alpha$" means "body".

1.2.3 Effects, diagnosis, treatment and fighting of schistosomiasis

The **effects on infected humans** are as follows: dermatitis and vomiting at the beginning, cough, temperature, pain in the chest (from eggs that get stuck in the blood vessels of the lungs), diarrhoea, blood in the urine, higher susceptibility to other diseases, kidney disease, cancer and cysts (especially in the liver).

Schistosomiasis is **diagnosed** mainly by detecting eggs, either in the sediment of urine or bladder biopsy or in the stool or through rectal biopsy depending on the type of schistosomiasis.

Treatment is with Praziquantel or (depending on the type of schistosomiasis) Metrifonate or Oxamniquine.

Measures against the spread of the disease are:

1) Treatment as mentioned above to kill the schistosomes in humans and therefore prevent eggs maintaining the cycle.

2) Sanitary installations and sewage treatment so that the eggs in the urine and faeces do not reach the area where there are intermediate hosts (snails).

3) Killing of snails through chemical substances or by supporting their natural enemies.

4) Preventing people from being infected by the cercariae through diverse infrastructural investments.

These measures have been implemented in various parts of the world. Some remarks should perhaps be made. The first measure is expensive in relation to the amount of money available in developing countries. But for tourists from rich countries who get infected it is a standard method. The second measure is surely very expensive and for the next few years not a high priority. Additionally it would be very impracticable in East Asia because looking at S. japonicum almost every mammal can serve as definitive host and the disease would be perpetuated in any case. This is too a main drawback of the first measure in East Asia. As regards the third measure, whilst the massive use of chemical substances to kill the snails created massive damage to the environment, the second possibility (support of the natural enemies of the snails) seems to be a good method, although it has not been possible to make the epidemic disappear completely. The fourth measure is used successfully too and does not necessarily need foreign capital to be implemented. Examples are the construction of simple bridges or sheltered sites where washing can be done. On the other hand it should be mentioned that other investments into the infrastructure, especially for irrigation have had a very bad impact on this epidemic and led to tremendous increases in the prevalence of this disease, that is the proportion of infected people, for instance in the Sahel.

The development of an effective vaccine has not yet been successful but research is continuing.

1.2.4 Social consequences and other diseases

Besides the tremendous direct health-problems associated with this disease and the consequences for the socio-economic situation there is a sex-specific point that should be mentioned. Women are infected more often with this disease and the infection usually turns out to cause more problems in women. The reason for women getting infected more often is that quite often in Africa *they* do the washing and fetch water and so are exposed more intensively. If a woman is infected, the symptoms are much the same as if she were infected by sexually transmissible diseases. So quite often women are worried to tell anybody about their pain and do not go to a medical doctor until the disease becomes chronic.

It might now be convenient to give the reader a brief overview of the importance of this disease in comparison to **other diseases**. In religious writings such as the bible we get quite a lot of information about diseases that bothered people **in the past**. As we have noted, bilharzia was such a disease. Other diseases, such as the plague, which were very important in former times, are no threat to the richer countries today and smallpox is an example of a disease which has been completely eradicated since about 1978. So the reader must be aware of the fact that the numbers below are **current numbers** and the situation was (especially for bilharzia) quite different 10, 30, 100 or 1,000 years ago.

The following summary does not take into account the severity of the diseases. The order of this summary is according to how many people are infected at present. Some diseases are mentioned, such as toxoplasmosis, which usually stay sub-clinical. On the other hand, other diseases which are clearly a threat to many countries and kill world-wide even more than 1 million people a year might not be mentioned. A description of these diseases and the impact can be found on the www-sites of the World Health Organisation (see below). Additionally we would point out that several worm-diseases are usually aggregated while we mention them all separately.

Toxoplasmosis	Perhaps about one third of the world's population are infected by toxoplasmosis. This is a disease caused by a protozoan which lives in cats, soil, and numerous mammals. People get infected by eating contaminated meat or vegetables.
Ascariasis	About 800 million people are infected with a nematode that causes ascariasis. Infection results from consumption of contaminated salad, vegetables, or water.
Hookworm	The hookworm-disease is caused by a nematode. About 670 million people are infected. Infection is by penetration through the skin.

Trachoma	Trachoma is a bacterial infection. About 600 to 700 million people are estimated to be infected with this disease. The infection can be transmitted simply by hand, towels or flies.
Pinworm	About 550 million people are infected by the pinworm. Infection results from uptake of eggs by carrier contacts, bed linen, food or dust.
Trichuriasis	Another nematode causes trichuriasis. About 520 million people are infected with this disease. People get infected by consumption of contaminated salads and vegetables.
Malaria	Malaria is one of the worst diseases and is caused by protozoans. About 300 to 500 million people get infected every year, killing up to one million annually. Transmission is by a mosquito.
Hepatitis	Hepatitis is caused by a virus. More than 2 billion people alive today have been infected with hepatitis B and some 350 million are chronically infected. Some 100 million are chronically infected with the hepatitis C virus. Depending on the type of hepatitis the infection is transmitted by consumption of contaminated food or water (hepatitis A) or direct or even indirect blood-contact (hepatitis B).
Giardiasis	About 200 million people suffer from giardiasis caused by a protozoan. Infection is by person-to-person transfer, sexual contacts or consumption of contaminated water or food.
Lymphatic filariases	About 120 million suffer from lymphatic filariases. This disease is caused by nematodes and transmission is by mosquitoes.
Strongyloidiasis	About 75 million people suffer from strongyloidiasis. This disease is caused by a nematode and infection is feco-oral.
Hymenolepiasis nana	A tapeworm causes hymenolepiasis nana. About 75 million people have this disease. Infection takes place via direct contact with other infected people or through ingestion of contaminated food or drinks.
Taeniasis saginata	About 40 to 60 million people suffer from taeniasis saginata. An infection with this tapeworm results from ingestion of smoked or undercooked beef such as tartare.
HIV/AIDS	About 33.4 million people suffer from HIV/AIDS and more than 12 million people have already died of this

disease (1998). Infection is mainly sexual. In 1995 more than 333 million new cases of **other sexually transmissible diseases (STDs)** occurred.

Further information about schistosomiasis can be found in Jordan and Webbe (1969, 1982), Stürchler (1988) and Basch (1991). Good articles (in German) about this disease can be found in the "Neue Zürcher Zeitung"; see Trüeb (1985) and Feldmeier (1995a, 1995b, 1997, 1998). Data on prevalence etc. of schistosomiasis is discussed in chapter 5 and compared with the simulated results using our models. Data on various diseases can be found at the following internet-sites:

1) "http://www.who.ch". This is the "World Health Organisation" in Geneva. The data above is from that site and Stürchler (1988).

2) "http://www.cdc.gov/cdc.html". This is the site of the "Center for Disease Control in Atlanta, USA".

1.3 Early models

We only give a very brief overview of earlier attempts to model this disease. For a more detailed discussion of the previous models we recommend the introduction to Kafetzaki's Ph.D. thesis (Kafetzaki (1993)) and Woolhouse (1991, 1992).

The first attempt to model the transmission of schistosomiasis was made by **Hairston** (1962, 1965). However, his model was not a dynamic model; he only looked at the equilibrium situation. It is not possible to make any predictions with his model.

The first dynamic model was developed by **George Macdonald**, at that time director of the Ross Institute of Tropical Hygiene in London, see Macdonald (1965). A slight simplification of his model was examined by Woolhouse (1992) and Barbour (1996). In their version, they did not distinguish between male and female worms but only looked at female worms that had mated. The main idea behind these models was that each living female parasite in a man produces a constant flow of infection into the snail-population throughout her life. On the other hand, each infected snail produces a constant flow of infection in the other direction. In these models, the factor limiting the development of the infection is that a snail produces the same flow of infection no matter how often it is infected. Such an assumption is justified through data of Pesigan et al (1958). Some limiting factor must be incorporated into a model in order to prevent infinite growth. This model does not represent reality in an acceptable way for two main reasons:

1) In typical models of epidemic spread there is a certain combination of parameters, the **"Basic Reproduction Ratio"**, which is central in determining the behaviour of the model. It is commonly denoted by R_0. R_0 is usually

defined to be the lifetime expected number of offspring of an adult parasite under ideal conditions. As can be easily imagined and as is usually the case, if $R_0 > 1$ an epidemic can develop and if $R_0 < 1$ it cannot develop. So eliminating the disease is equivalent in these terms to altering the relevant parameters so that R_0 becomes smaller than 1. This is a theoretical concept. In the models of Macdonald (1965), Woolhouse (1992) and Barbour (1996, pages 136 and 137), R_0 can apparently be estimated in the field merely by measuring the prevalence of infection in snails: In these models, if $\overline{y}$ denotes the equilibrium proportion of infected snails, the relationship between R_0 and $\overline{y}$ is

$$R_0 = \frac{1}{1-\overline{y}}.$$

The levels of prevalence of infection in snails found in practice are usually very low, of the order of 0.1-10%, roughly corresponding to values of R_0 between 1 and 1.1. A range so narrow would be astonishing in view of the variability of the parameters involved in the definition of R_0 from place to place, and it is unlikely that R_0 should everywhere have a value that was just enough to sustain transmission, and no more, especially in view of the apparent resistance of transmission to attempts at control. So something must be wrong.

2) In Barbour (1978) a period of latent infection in the snails was included. Additionally a seasonal and spatial heterogeneity amongst the human population was studied. The conclusion was that without including a non-linear mechanism *in the definitive host* , the model cannot mirror real data appropriately. Therefore immunity (or crowding effects) in the human population must be included.

Other models (stochastic and deterministic) have other drawbacks such as:

3) The distribution of the number of parasites per human: Looking at the distribution of the data of Cheever (1968), the ratio "variance / mean" was 537, while in many models, due to the assumptions used in building these models, this ratio was 1 (Poisson distribution of parasites per human).

The reader has seen that there have been a variety of attempts to model this disease. Therefore we can ask ourselves: "which model should be chosen?" In Barbour (1996) this question is answered in the following way:

> A model is valuable insofar as it enables one to better understand the phenomenon. No model can be expected to mirror every aspect precisely, and it may well be the case that a variety of models are equally valid, in the sense that they give useful information about different aspects. The more precise the information desired, the more detailed the model has to be; however, the more detailed the model, the greater the danger that it becomes incomprehensible. A reasonable approach consists of both simple and detailed models, the detailed models for quantitative predictions, and

the simple models for qualitative understanding and for checking the detailed model. None of the models should be made more detailed than is necessary for the purposes for which they are used.

1.4 Our models and notation

In the above title the word "our" stands for the following researchers: Prof. Dr. A.D. Barbour, Dr. Maria Kafetzaki, Dr. J.A.P. Heesterbeek and myself.

In total eight models will be developed: stochastic and deterministic, linear and non-linear, with and without mortality of humans; this explains why we have 8 models: $2 \times 2 \times 2 = 8$. As mentioned above, the four stochastic and the four deterministic models are connected to each other in a one to one fashion. The 4 linear models are also connected to the 4 non-linear models one to one in the sense that each linear model serves as a good (and mathematically much easier to handle) approximation to the initial phase of the corresponding non-linear model. In what follows we first discuss the assumptions we are going to make. The assumptions are almost the same for all 8 models because the 8 models are closely related. Later we write down the 8 models, giving the rates of change in the stochastic models and the systems of differential equations in the deterministic models.

The assumptions we make for all models are as follows. We suppose that people make potentially infectious contacts according to a Poisson process of rate λ. These people act independently of one another and "mix infinitely fast", meaning that if such a contact occurs at any time t, the infectious person meets any other person with the same chance. If an infected person makes a potentially infectious contact with another individual who is already infected nothing happens, so there is no super infection, that is we assume concomitant immunity. But in the linear models we look at the initial phase of the epidemic and therefore assume that every contact is made with an uninfected individual. These assumptions are translated into mathematical language in the following way: in the non-linear models we have to alter the contact-rate λ to be $\lambda\xi_0$ and λx_0 respectively (where ξ_0 and x_0 represent the *proportions* of uninfected people in the deterministic and stochastic models respectively) while in the linear models we can leave the contact-rate λ unchanged.

It might be necessary to point out a misunderstanding that might emerge. We have said that people act independently of one another in infecting other people and the infection process is therefore modelled with independent Poisson processes. But if somebody infects someone else, there are always two people involved. So one might think that the Poisson processes are not independent of each other. We therefore make clear that (in correspondance with the biological situation) there is always a person that infects, and a person that (at least potentially) gets infected. The infection process of every person is modelled with a Poisson process and then we can choose these infection processes independently of each other.

In contrast to most other models, none of the eight models directly includes the intermediate host. We make the following assumptions. If person A having k worms has a potentially infectious contact with an uninfected person B, then an infection with j worms develops in person B with probability p_{kj}. Clearly we have $\sum_{j\geq 0} p_{kj} = 1$ for all k and we additionally assume that $p_{00} = 1$. We denote by F_k the distribution $\{p_{kj}, j \geq 0, j \in \mathbb{N}\}$. Since we assume that the worms act independently of one another, F_k is the $k-$th convolution of F_1. We denote the mean of F_1 by $\theta > 0$, that is $\theta := \sum_{j\geq 1} j p_{1j}$ and the variance by $\sigma^2 < \infty$, that is $\sigma^2 := \sum_{j\geq 0}(j-\theta)^2 p_{1j}$. The convolution F_k has mean $k\theta$ and variance $k\sigma^2$.

If a person gets infected he carries one or more worms in his body. We assume that these are all female worms or better: we do not distinguish at all amongst the sexes and assume therefore that a single worm can produce eggs. We assume that these worms act independently of one another in producing eggs and in their length of life (and therefore in the length of time they produce eggs). We model the length of life of a worm in the human body with an exponential distribution with parameter μ.

In the models with mortality of humans we model the length of life of each person with an exponential distribution with parameter κ. We assume that an infection does not increase mortality of humans and people die independently of each other. Clearly all worms die if their host dies, moreover, in the non-linear models we assume that as soon as a person dies he is replaced by an uninfected person (a child maybe).

Some remarks should be made about whether these assumptions are justified.

The assumptions of exponentially distributed lifetimes of worms in humans and of humans themselves are mainly a compromise to make the mathematics easier.

Then we implicitly make the assumption that the infection process is by group-infection and not trickle-infection: We assume that people get all their burden of infection at once (the j worms from above), and not one at a time. This is now a clear assumption which is not easy to prove in real life. The theoretical thoughts behind this theory are that eggs are deposited in large numbers in an area and so lots of snails are infected there. So, when a person enters that area he risks being infected by several cercariae. In particular it is a matter of chance (or better bad luck) if someone gets infected at all and if so with how many worms. This theory has been developed in Barbour (1977).

An important assumption is concomitant immunity as a limiting factor in human beings. Concomitant immunity means that people get infected and then are immune to further infections until they have fully recovered. There are immunological theories which support this assumption, but the mechanism is not fully understood yet. Looking at real data, the situation seems very

complex: In Pesigan et al. (1958) the prevalence of schistosomiasis japonicum was studied. Looking at the proportion of infected as a function of age we have an increase until the age of about 15 years and then a levelling off. This could be due to concomitant immunity as we model it. In fact, looking at simulated results in chapter 5.3, we get very much the same picture through simulations as we have in real data. Looking at other types of schistosomiasis such as S. mansoni we have the following situation: If we look at the proportion of infected people as a function of their age (see Smithers and Terry (1969), Warren (1973), Kloetzel and da Silva (1967)) we see that the proportion of infected people increases dramatically until the age of about 15 and then decreases again until the age of 30 where it stabilises. People might argue that the contact-rate (λ) is much smaller in that age-group because older people do not have so much water-contact as young children. But as Klötzel and da Silva (1967) showed, looking at immigrants who come into an endemic region, that it is the time that a person spends in such a region that lets him develop an acquired immunity. Our particular model of concomitant immunity does not lead to the decline from ages 15 to 30 described above. Another sort of (acquired concomitant) immunity is suggested in section "5.4.2.3 Concomitant immunity after 2 infections ACQUIRED".

Researchers are uncertain about what is really going on. Mathematicians merely assume a theory and then look whether predicted data from simulations based on models fit real data well. If not, the assumptions are not justified; if yes, it might well be that there is something going on in reality which behaves as postulated in our models.

So now all eight models can be written out in terms of rates at which the state changes in the stochastic models or as systems of differential equations in the deterministic models. First we formulate the stochastic models.

The first stochastic model was developed in Barbour and Kafetzaki (1993). Suppose that there are M individuals in our system. Let $x^{(M,0)}$ be an infinite dimensional Markov process

$$x^{(M,0)}(\omega, t) : \Omega \times [0, \infty) \to \{[0, 1] \cap M^{-1}\mathbb{Z}\}^{\infty}.$$

We assume that $\sum_{j\geq 0} x_j^{(M,0)}(0) = 1$ and $x_j^{(M,0)}(0) \geq 0$, $j \geq 0$. In this model $x_j^{(M,0)}(t)$, $j \geq 0$, denotes the *proportion* of individuals at time t, $t \geq 0$, who are infected with j worms. We suppress the index M in what follows. The rates at which $x^{(0)}$ changes are as follows (superscript (0) and time t suppressed too):

$$\begin{aligned} &x \to x + M^{-1}(e_{j-1} - e_j) \text{ at rate } jM\mu x_j \ ; \ j \geq 1, \\ &x \to x + M^{-1}(e_k - e_0) \text{ at rate } \lambda M x_0 \sum_{l\geq 1} x_l p_{lk} \ ; \ k \geq 1, \end{aligned} \tag{1.3}$$

where e_i denotes the i-th co-ordinate vector in $\mathbb{R}^\infty$ in the whole thesis. We call this model SN; this stands for **S**tochastic **N**on-linear. We introduce a notation for the sigma-algebra too: $\mathcal{F}_s := \sigma\{x^{(0)}(u), 0 \leq u \leq s\}$. At this point it is convenient to explain why we have *these* rates: there are jMx_j worms in individuals with j worms and they all die at a rate of μ. If such a worm dies, the proportion of individuals with j worms decreases by $1/M$ and the proportion of individuals with $(j-1)$ worms increases by $1/M$. This explains the first transition-rate. The second transition is an infection: there are $x_l M$ individuals with l worms who make contacts according to a Poisson process of rate λ. But only those contacts that take place with uninfected individuals are infective. So the rate must be decreased by multiplying by the proportion of uninfected people x_0. Then we must include the probability that such an infection leads to an infection with k worms, hence the probability p_{lk}. All other rates are explained in a similar way.

The next stochastic model was developed in Barbour (1994). Let $X^{(0)}$ be an infinite dimensional Markov process

$$X^{(0)}(\omega, t) : \Omega \times [0, \infty) \to \{[0, \infty) \cap \mathbb{Z}\}^\infty.$$

We assume that $0 < \sum_{j \geq 1} X_j^{(0)}(0) = M < \infty$ and $X_j^{(0)}(0) \geq 0$, $j \geq 1$. In this model $X_j^{(0)}(t)$, $j \geq 1$, denotes the *number* of individuals at time t, $t \geq 0$, that are infected with j worms. The rates with which $X^{(0)}$ changes are as follows (superscript (0) and time t suppressed):

$$\begin{aligned} &X \to X + (e_{j-1} - e_j) \text{ at rate } j\mu X_j \ ; \ j \geq 2, \\ &X \to X - e_1 \text{ at rate } \mu X_1 \ ; \ (j = 1), \\ &X \to X + e_k \text{ at rate } \lambda \sum_{l \geq 1} X_l p_{lk} \ ; \ k \geq 1. \end{aligned} \tag{1.4}$$

We call this model SL; this stands for **S**tochastic **L**inear. We introduce a notation for the sigma-algebra too: $\mathcal{G}_s := \sigma\{X^{(0)}(u), 0 \leq u \leq s\}$. The difference between models SN and SL is the following: in model SL the contact rate is λ and there is no limiting factor in the model. In model SN the contact rate is altered from λ to $\lambda x_0^{(0)}$, because only those infectious contacts that are with an uninfected individual lead to a new infection.

The next two stochastic models include mortality of humans. Therefore we no longer attach the superscript (0) in the mathematical symbols x and X respectively. Let $x^{(M)}$ be an infinite dimensional Markov process

$$x^{(M)}(\omega, t) : \Omega \times [0, \infty) \to \{[0, 1] \cap M^{-1}\mathbb{Z}\}^\infty.$$

Note that the index M is suppressed in what follows. In this model we assume that $\sum_{j \geq 0} x_j(0) = 1$ and $x_j(0) \geq 0$, $j \geq 0$. $x_j(t)$, $j \geq 0$, denotes the *proportion*

of individuals at time t, $t \geq 0$, that are infected with j worms. The rates with which x changes are as follows (time t suppressed too):

$$\begin{aligned} &x \to x + M^{-1}(e_{j-1} - e_j) \text{ at rate } jM\mu x_j \ ; \ j \geq 1, \\ &x \to x + M^{-1}(e_k - e_0) \text{ at rate } \lambda M x_0 \sum_{l \geq 1} x_l p_{lk} \ ; \ k \geq 1, \\ &x \to x + M^{-1}(e_0 - e_r) \text{ at rate } M x_r \kappa \ ; \ r \geq 1. \end{aligned} \tag{1.5}$$

We call this model SNM; this stands for **S**tochastic **N**on-linear including **M**ortality of humans. We introduce a notation for the sigma-algebra too: $\mathcal{H}_s := \sigma\{x(u), 0 \leq u \leq s\}$.

The last stochastic model is defined as follows. Let X be an infinite dimensional Markov process

$$X(\omega, t) : \Omega \times [0, \infty) \to \{[0, \infty) \cap \mathbb{Z}\}^\infty.$$

We assume that $0 < \sum_{j \geq 1} X_j(0) = M < \infty$ and $X_j(0) \geq 0$, $j \geq 1$. In this model $X_j(t)$, $j \geq 1$, denotes the *number* of individuals at time t, $t \geq 0$, that are infected with j worms. The rates at which X changes are as follows (time t suppressed):

$$\begin{aligned} &X \to X + (e_{j-1} - e_j) \text{ at rate } j\mu X_j \ ; \ j \geq 2, \\ &X \to X - e_1 \text{ at rate } \mu X_1 \ ; \ (j = 1), \\ &X \to X + e_k \text{ at rate } \lambda \sum_{l \geq 1} X_l p_{lk} \ ; \ k \geq 1, \\ &X \to X - e_r \text{ at rate } X_r \kappa \ ; \ r \geq 1. \end{aligned} \tag{1.6}$$

We call this model SLM; this stands for **S**tochastic **L**inear including **M**ortality of humans. We introduce a notation for the sigma-algebra too: $\mathcal{I}_s := \sigma\{X(u), 0 \leq u \leq s\}$. The difference between model SNM and SLM is the following: in model SLM the contact rate is λ and there is no limiting factor in the model. In model SNM the contact rate is altered from λ to λx_0, because only those infectious contacts that are with an uninfected individual lead to a new infection.

In our assumptions we have a Poissonian contact-process with parameter λ and all lifetimes are exponentially distributed (with parameters μ and κ respectively). So it was obvious how the *stochastic* models had to be built, that is how the transition rates had to be. But in *deterministic* models this is not obvious. So we might just take the four deterministic models below as a proposal of models that mirror the assumptions above appropriately. But later, one of the things we prove (in chapter 3) is that in fact the deterministic models are the (weak) limits of the stochastic models.

The first deterministic model is that developed in Barbour and Kafetzaki (1993): Let $\xi_j^{(0)}(t),\ j \geq 0$, denote the *proportion* of individuals at time t, $t \geq 0$, who are infected with j worms. We suppose that $\sum_{j\geq 0} \xi_j^{(0)}(0) = 1$ (whereas the equation $\sum_{j\geq 0} \xi_j^{(0)}(t) = 1$ for all $t \geq 0$ had to be proven in Barbour and Kafetzaki (1993)). We assume that $\xi_j^{(0)}(0) \geq 0$ for all $j \geq 0$. $\xi^{(0)}$ is the solution of the following system of differential equations (superscript (0) and time t suppressed):

$$\begin{aligned} \frac{d\xi_j}{dt} &= (j+1)\mu\xi_{j+1} - j\mu\xi_j + \lambda\xi_0 \sum_{l\geq 1} \xi_l p_{lj};\ j \geq 1, \\ \frac{d\xi_0}{dt} &= \mu\xi_1 - \lambda\xi_0(1 - \sum_{l\geq 0} \xi_l p_{l0}). \end{aligned} \tag{1.7}$$

We call this model DN; this stands for **D**eterministic **N**on-linear.

The next model was first studied in Barbour, Heesterbeek and Luchsinger (1996). Let $\Xi_j^{(0)}(t),\ j \geq 1$, denote the *number* of individuals at time t, $t \geq 0$, who are infected with j worms. We assume that $\Xi_j^{(0)}(0) \geq 0$ for all $j \geq 1$. $\Xi^{(0)}$ is the solution of the following system of differential equations (superscript (0) and time t suppressed):

$$\frac{d\Xi_j}{dt} = (j+1)\mu\Xi_{j+1} - j\mu\Xi_j + \lambda \sum_{l\geq 1} \Xi_l p_{lj};\ j \geq 1. \tag{1.8}$$

We call this model DL; this stands for **D**eterministic **L**inear. The difference between model DN and DL is the following: in model DL the contact rate is λ and there is no limiting factor in the model. In model DN the contact rate is altered from λ to $\lambda\xi_0^{(0)}$, because only those infectious contacts that are with an uninfected individual lead to a new infection. There is an important point that has to be mentioned in view of model DL: We said that $\Xi_j^{(0)}(t)$ denotes the *number* of individuals at time t who are infected with j worms. But $\Xi_j^{(0)}(t)$ is a solution to a differential equation and therefore this can be any real number, and in particular very small. Then it is not very convenient to say that $\Xi_j^{(0)}(t)$ denotes the number of people with j parasites. $\Xi_j^{(0)}(t)$ is rather the *expected* number of people with j parasites. In fact the linear model DL is mainly used to model a situation where there is no limiting factor (no non-linearity). The number of infected people (in reality) should be small *in comparison* to the number of people in the system (in model DL we even assume that there are infinitely many susceptibles). But as soon as the number of infected people gets small not only in comparison to the number of people in the system but

absolutely, the model is not appropriate to mirror reality. The same is true for model DLM (see below for a definition).

The next two deterministic models are the same as above but with mortality included. Let $\xi_j(t),\ j \geq 0$, denote the *proportion* of individuals at time $t,\ t \geq 0$, who are infected with j worms. We suppose that $\sum_{j\geq 0} \xi_j(0) = 1$ (whereas the equation $\sum_{j\geq 0} \xi_j(t) = 1$ for all $t \geq 0$ has to be proven). We assume that $\xi_j(0) \geq 0$ for all $j \geq 0$. ξ is the solution of the following system of differential equations (time t suppressed):

$$\begin{aligned} \frac{d\xi_j}{dt} &= (j+1)\mu\xi_{j+1} - j\mu\xi_j + \lambda\xi_0 \sum_{l\geq 1} \xi_l p_{lj} - \kappa\xi_j;\ j \geq 1, \\ \frac{d\xi_0}{dt} &= \mu\xi_1 - \lambda\xi_0(1 - \sum_{l\geq 0} \xi_l p_{l0}) + \kappa(1-\xi_0). \end{aligned} \tag{1.9}$$

We call this model DNM; this stands for **D**eterministic **N**on-linear including **M**ortality of humans.

The next is just the linear version of the previous model. Let $\Xi_j(t),\ j \geq 1$, denote the *number* of individuals at time $t,\ t \geq 0$, who are infected with j worms. We assume that $\Xi_j(0) \geq 0$ for all $j \geq 1$. Ξ is the solution of the following system of differential equations (time t suppressed):

$$\frac{d\Xi_j}{dt} = (j+1)\mu\Xi_{j+1} - j\mu\Xi_j + \lambda \sum_{l\geq 1} \Xi_l p_{lj} - \kappa\Xi_j;\ j \geq 1. \tag{1.10}$$

We call this model DLM; this stands for **D**eterministic **L**inear including **M**ortality of humans. The difference between model DNM and DLM is the following: in model DLM the contact rate is λ and there is no limiting factor in the model. In model DNM the contact rate is altered from λ to $\lambda\xi_0$, because only those infectious contacts that are with an uninfected individual lead to a new infection.

In both deterministic linear models, that is DL and DLM, we restrict our studies to solutions Ξ for which the following three conditions are satisfied:

$$\begin{aligned} &\sup_{0\leq s\leq t} \sum_{j\geq 1} \Xi_j(s) < \infty \ \text{ for all } t \geq 0, \\ &\sum_{j\geq 1} j\Xi_j(0) < \infty, \\ &\text{there exists a } j \geq 1, \text{ such that } \Xi_j(0) > 0. \end{aligned} \tag{1.11}$$

We call these constraints conditions C. They are introduced for technical reasons. In the following chapters the initial values are at times more general than specified above (but if so, we shall make it clear).

We only consider nonnegative solutions to the deterministic differential equations, even if we do not mention it every time.

To end this part let us repeat the notations introduced so far: The **mathematical symbol** for a stochastic model is a small x or a capital X all the way through and for the deterministic model it is a small ξ or a capital Ξ. The non-linear models have a small letter (x or ξ) and the linear models a capital letter (X or Ξ). In the linear models the symbol means *number* of individuals, while in the non-linear models it means *proportion* of individuals. But then a superscript (0) is attached if the mortality rate κ is taken to be zero. So the symbols for all eight models are as follows:

Specification (Abbrev.)	Equation	Symbol
Stoch. non-linear without mortality (SN)	(1.3)	$x^{(0)}$
Stoch. linear without mortality (SL)	(1.4)	$X^{(0)}$
Stoch. non-linear with mortality (SNM)	(1.5)	x
Stoch. linear with mortality (SLM)	(1.6)	X
Det. non-linear without mortality (DN)	(1.7)	$\xi^{(0)}$
Det. linear without mortality (DL)	(1.8)	$\Xi^{(0)}$
Det. non-linear with mortality (DNM)	(1.9)	ξ
Det. linear with mortality (DLM)	(1.10)	Ξ

Table 1.1: Table of symbols used for the eight models.

The **symbols for the four sigma-algebras** in the stochastic models are as follows:

$\mathcal{F}_s := \sigma\{x^{(0)}(u), 0 \le u \le s\}$.
$\mathcal{G}_s := \sigma\{X^{(0)}(u), 0 \le u \le s\}$.
$\mathcal{H}_s := \sigma\{x(u), 0 \le u \le s\}$.
$\mathcal{I}_s := \sigma\{X(u), 0 \le u \le s\}$.

The **basic reproduction ratios** "R_0" have to be distinguished too. But here the situation is easier because they are the same, no matter whether the system is stochastic or deterministic, linear or non-linear, as is to be shown. The differences arise because of changes in θ; there are two relevant regions for θ in the case without mortality of humans and three in the case with mortality of humans. We denote those ratios by $R_0^{(0)}, R_1^{(0)}$ and R_0, R_1, R_2 respectively.

The reader might ask himself why we look at as many as *eight* models. The reasons are as follows:

1) We want to look at stochastic and deterministic models because they each have their advantages and disadvantages as discussed in 1.1. The stochastic models are closer to reality but more complex to analyse; the deterministic

models mirror only an "average" behaviour of the process, but we can analyse them more easily.

2) The linear models are easier to handle than the non-linear models but they are only useful for the initial phase (because in the linear models we assume that *every* contact is with an uninfected person).

3) The introduction of mortality of humans is an attempt to model reality in a better way. This point is discussed in chapter 5.

4) Additionally, for those interested in the proofs, the eight models are linked to each other in various ways which enabled us to prove a large variety of theorems in one model by comparison with analogous results in another model.

Remark: To end this chapter we might add that our models could be considered for modelling the hookworm disease too. In this case the description is even closer to reality, because no intermediate host exists in the transmission of that disease.