Kees van Heugten, Servaas Houben and Rob Smit

Modelling COVID, using simulation techniques, looking at things from a different perspective

The Covid-19 pandemic means studying models that describe virus behaviour have become of interest to the community and the general public. During the pandemic, governments introduced metrics like the reproduction factor, the number of infected people, the number of people hospitalised etc. as steering information. The main objective for governments is to control the virus (flatten the curve) and to be able to have sufficient hospital and ICU capacity for people who have become severely ill because of the virus. Currently not many countries are managing this capacity issue well. In the Netherlands, a new partial lockdown started mid October to try to lower the number of Covid-19 patients since not only Covid-19 patients are using ICU beds but also regular patients too. On average at least 70% of hospital bed capacity is being used in the Netherlands, so adding extra Covid-19 patients can lead to capacity problems. And this is now what is happening around the world.

Publicly available models, like the SIR model, use a theoretical framework to investigate the spread of the virus within a community. So, they tend to project only the most likely outcomes in terms of the number of susceptible, infected and recovered people. These models use data from the past and in general do not make a differentiation in terms of age or asymmetric infection behaviour, nor do they include the different states of the virus. Since the data is, in our view, still not stable enough and age is a deciding factor as is the way the virus spreads, we decided to use a different approach to simulating the Covid-19 virus. First we defined a Markov framework and model, then we used this to model the spread of the virus with a stochastic approach adding age as a differentiation indicator. Moreover, we added other elements that will be discussed in this document.

The missing element in the government dashboards is being able to define metrics to predict future hospital and ICU capacity needs and confidence intervals in order to be able to act before it is too late; in other words, an early warning indicator. In this article, a model containing the hospital and ICU capacity prediction is introduced, including confidence intervals around it. This model has been used as a dashboard in the Netherlands since 15 September 2020 and predicts hospital and ICU capacity 28 days ahead. The dashboard is published every week on www.crownactuaries.eu.

The results so far look very promising. Our last estimation of Covid-19 patients from September 22 in hospital beds lies within our 95% confidence interval. Our current modelling approach is universal and could also be used in other countries.

The status of an individual in the population

In general, the virus model makes use of the status an individual can have in relation to the virus. An individual can be healthy (never infected), can be infected, will become sick (mildly, medium or severely), can recover from sickness and be immune for a certain period, can become healthy again or die from the sickness. The reproduction factor is the driver of the infection process. This factor is assumed to be time dependent. The base reproduction factor $R0$ is defined as the number of people who will be infected in a given time period by one infected person if the virus is not controlled. For Covid-19 $R0$ is around 2.5. Note that $R0$ is not a biological constant since it will be affected by other factors such as environmental conditions and the behaviour of the population. If we allow the virus to go uncontrolled the effective reproduction factor will be lower than the base factor due to the fact that, in the population, there will be people who are already infected, sick, immune or have died. As a result, the maximum number of people who can be infected will be less than the total population. The process of the change in status for the individuals in the population is described as a Markov model. Depending on the time for being sick

Figure 1:
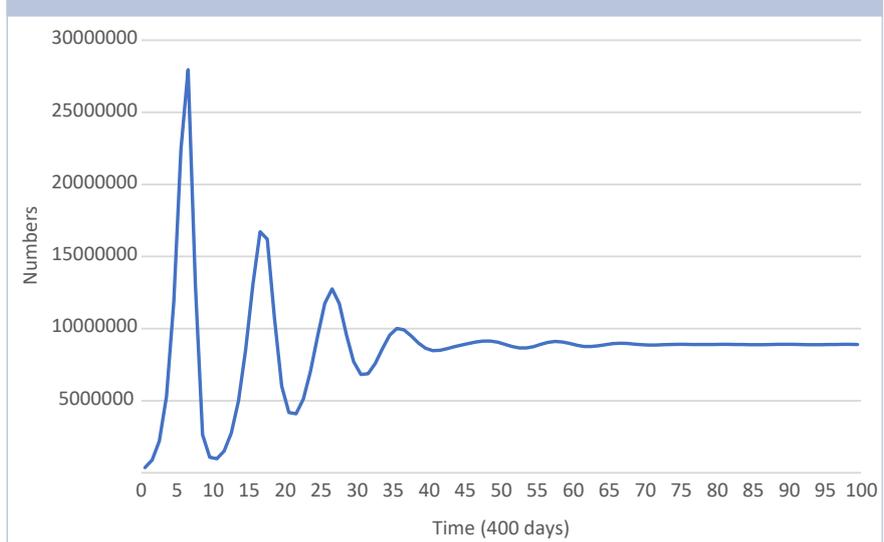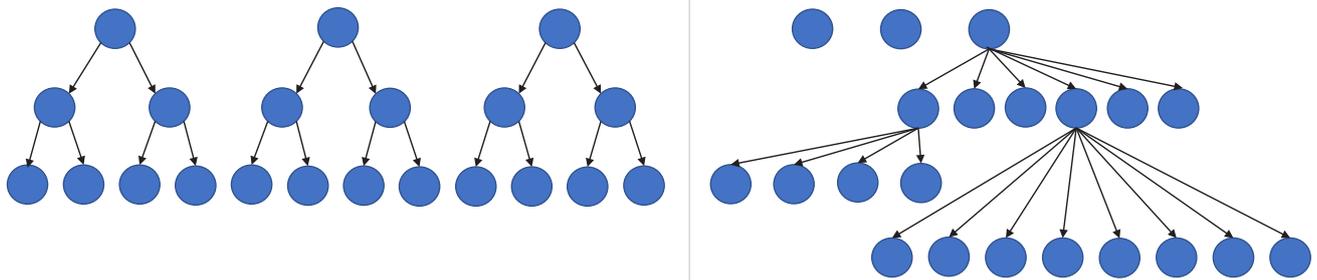**Development of the number of infections for Germany uncontrolled**

Figure 2:
Symmetrical and asymmetrical infection process

The picture on the left shows a symmetrical infection process, with reproduction factor 2. Every infected person infects two others in the next period. In the second period, the six infected persons infect two others again. The picture on the right shows an asymmetrical infection process, with reproduction factor 2. Of the first three infected only one infects six others, while the other two don't infect anyone. In the second period, only two infect while the other four don't, Of the two that infect, the first infects four others and the second infects eight others. In total the reproduction factor is the same, although the infection process at a micro level is different.

and immune the Markov process can result in a steady state situation, with a fixed number of individuals being healthy, infected, sick or immune. To arrive at this steady state situation one has to exclude the status deceased otherwise the steady state situation is trivial, with all individuals being deceased.

Figure 1 shows the development of the number of infected assuming that the virus is not controlled, starting with 360,000 confirmed infections (as reported in Germany for 18 October) on a population of 87.8 million. We assume that nobody dies, and there are currently no sick or immune people. The assumptions regarding the duration of sickness and immunity can be found on the website (under assumptions). We observe that the number of infections fluctuates and finally stabilizes at 10.6% (8.9 million) of the population. There will be 4 million permanently severely ill people in the steady state. The number of healthy people is 33.5 million, being 40% of the total population, with the 40% being 1/2.5 (1/R0).

The asymmetrical behaviour of Covid-19

Figure 2 shows the difference between a symmetrical and asymmetrical infection process. In a symmetrical infection process, every infected person is assumed to infect others in line with the reproduction factor. For an asymmetrical infection process

there are infected people who don't infect others (or less than is expected according to the reproduction factor) and there are others who infect many more other people than is assumed by the reproduction factor. This aspect is described by Smith in Nature (2005). Due to this behaviour we see local outbreaks of the virus and 'super spread events'.

Deterministic models that work with a population reproduction factor do not take this behaviour into account, which can result in an under estimation of the expected capacity for hospitals and ICUs, especially locally. In our Covid-19 model we therefore introduce individual reproduction factors that are assumed to be drawn from a skewed probability

distribution. To get the required skewness we use a discrete Gamma distribution, where we take into account the fact that we need to align with the total reproduction factor at the population level and where we can change the parameters to get the required skewness. In the following case study we simulate on a population of 10,000 and use 4.33 and 0.3 as gamma parameters (the population R was 1.3 at that moment). Gamma distribution is a well known instrument used in Non-Life actuarial models to estimate, for example, average claims for car insurance. The parameters are determined based on expert judgement. The skewness of the gamma distribution will depend on the measures that are taken such as forbidding

Figure 3:
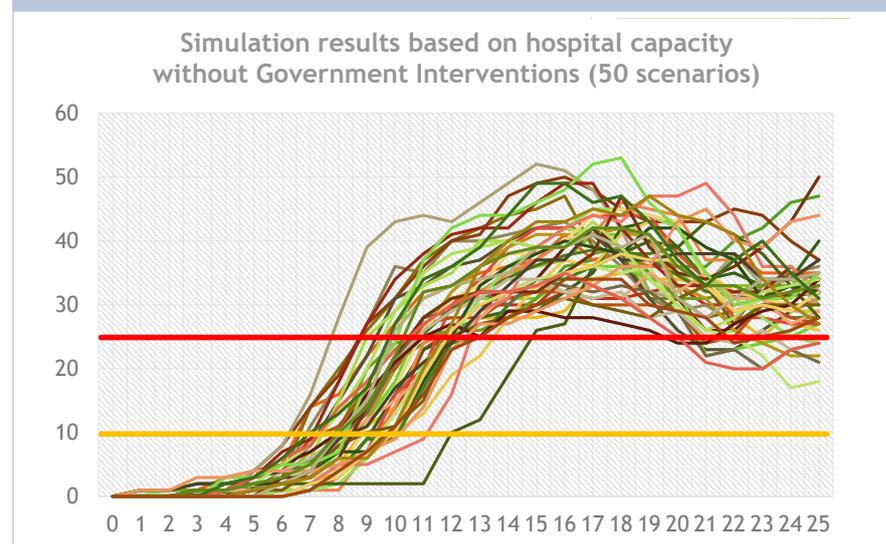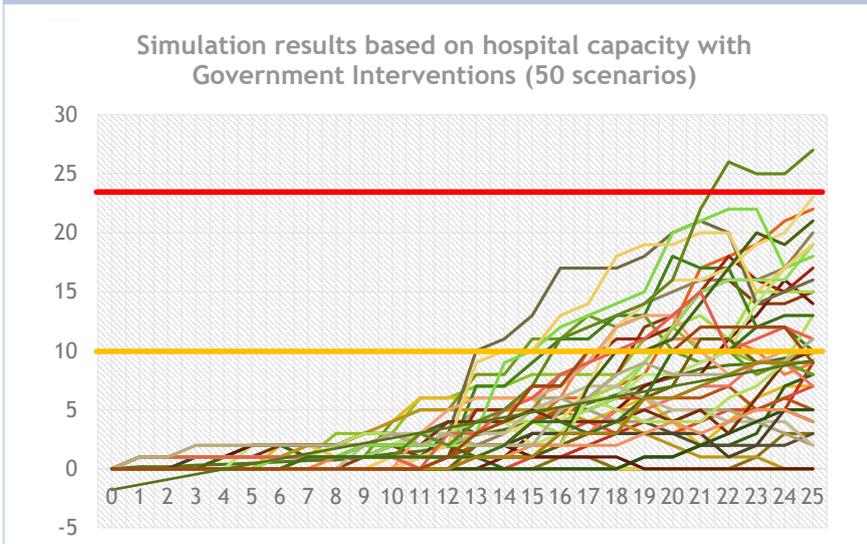**Number of severely sick, without intervention**

Simulation results based on hospital capacity without Government Interventions (50 scenarios)

Figure 4:
Number of severely sick, with intervention

Simulation results based on hospital capacity with Government Interventions (50 scenarios)

mass events and limiting group sizes.

Simulations using the asymmetrical behaviour of Covid-19

In figure 3 we see the effect of 50 scenarios based on a population of 10,000 and 4 people infected (this was in line with the situation in the Netherlands in July 2020, during the first wave.

The maximum hospital capacity is 23 beds per 10,000 (around 40,000 beds for the total population in the Netherlands). In these simulations it is assumed that there are 10 beds

available for Covid-19 patients, meaning around 17,500 beds, and that no government interventions are taken. The reproduction factor is 1.3 in four days. One can see that the average number of beds grows until it reaches 40 per 10,000, almost twice the number of beds available.

Due to the stochastic nature there are even scenarios in which the number of beds needed is more than 50 per 10,000. It is also clear that there are no scenarios that remain below the maximum capacity of 23. This simulation shows that no intervention is not an option if the government wants to assure medical care for Covid-19 patients as well as for others.

Figure 4 illustrates the number of severely sick individuals per 10,000 inhabitants, with government intervention. When the number of severely sick individuals is above 2 per 10,000 or the number of infected individuals is more that 1% of the population, measures are introduced to reduce the reproduction factor to 1. As a result, one can observe that on average the number of severely sick individuals has a maximum value of 10 per 10,000 which matches the capacity of available beds. Looking at the 50 scenarios that are simulated we see that planning based on the average number can still result in ca-

pacity problems in relation to the number of available beds: there are scenarios that require almost 50% extra capacity than expected. When planning the capacity, one has to take into account the capricious nature of the Covid-19 infection process.

Introducing a stochastic dashboard

As mentioned, the current dashboards used by the governments do not predict the number of beds, although it is clear that there is a need to know how many Covid-19 beds will be needed in the near future. Based on our model, we can predict the number of beds required in the coming period, taking into account the fact that the government will act if the number of infections grows too quickly. The government will act on its dashboard and with our model we can calculate the effect of the government interventions. We produced a barometer that predicts the number of hospital beds and ICUs for Covid-19 patients within a horizon of 28 days. Our first prediction was made on 22 September for October 20:

We expected an average number of hospital and ICU beds of 1,549, and up to 1,888 with a confidence of 95%, as shown in figure 5. The actual number of hospital and ICU beds tur-

Figure 5:
Prediction of the number of hospital beds and ICUs for 20 October (Yellow bar: 20% of the total reflects the number of ICUs)

Covid-19 Barometer for the Netherlands including IC capacity
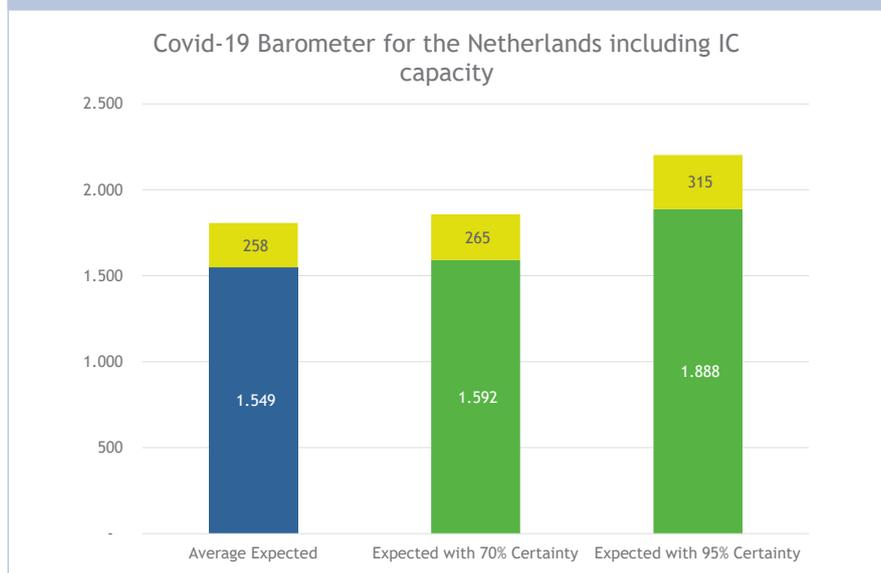
ned out to be 1,859, i.e., within the 95% confidence range.

With this barometer it is easier for policymakers to make timely decisions to avoid additional measures being required. All the information from the current dashboards is translated in the number of expected hospital beds and ICUs.

Calibrating the Covid-19 barometer

Since the Covid-19 pandemic a lot of information has become available but a lot remains unknown. In our virus model the number of infections is an important driver of the number of severely sick persons at a later stage. But the number of infected is not known, since we don't test all individuals in the population. Also testing does not guarantee that an individual is infected if they test positive. In our assumptions we assume a certain percentage of the infected will become severe ill due to Covid-19 depending on the age of the individual. We are well aware of the fact that our model needs to be calibrated based on the data used to give correct predictions. Our approach is that we

Figure 6:
The model
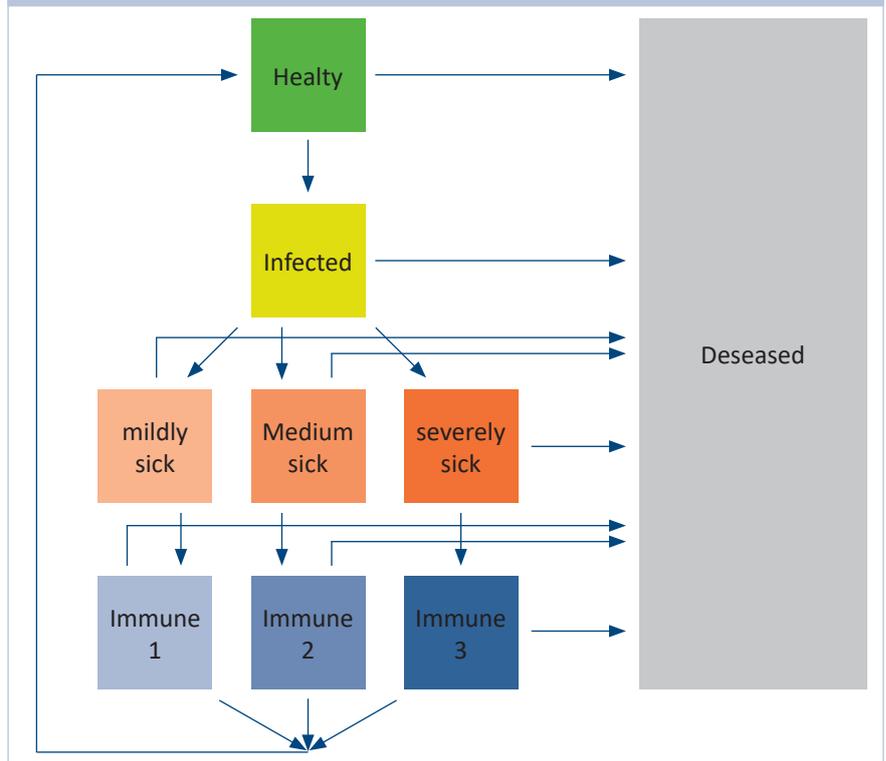
backtest the numbers of hospital and ICU beds we predicted compared to the reality. Based on the difference between the prediction and reality we will change our assumptions.

Currently we have a prediction for 12 November 2020. Since we publish the barometer every week, we will gather enough information to calibrate our model.

How the model exactly works

There are two models that predict the development of the virus. The first one is a population model, that divides the population into groups:

• Healthy
• Infected
• Sick (mildly, middle, severely)
• Immune (short, middle, long)
• Deseased

The relations between the statuses are illustrated in the following diagram

The infection process is modelled as (from time t to t+1):

Number of new infected(t+1) = number of existing infected(t) * R0 * Number of healthy(t) / Total population number(t).

Where the expression R0*Number of healthy(t)/Total population number(t) is the effective reproduction factor R(t). Note that the number of new infected is maximized to the number of healthy(t) and the number of healthy individuals who die in the period [t, t+1].

After being infected and passing the incubation time the group will become

sick: mild (asymptomatic), middle (symptomatic) and severe (in hospital or ICU). The chances of becoming mildly, middle or severely sick are age dependent. On average 80% will become mildly sick, 15% middle and 5% severely, but this depends on the age structure of the population. The population model works with the average chances. After being sick the group will become immune. It is expected that the immune time is dependent on the level of sickness, where a mildly sick person is immune for only a short period, a middle sick person for a little longer and a severely sick person for the longest period. After the immune period the group is healthy again. It is assumed that this healthy group is the same as never infected.

The chances of death are modelled for every status and are age- and status-dependent. In our calculations different (higher) mortality rates are used for the severely sick.

The population model can be found on our website as an educational tool in excel. One can also assume that asymptomatic groups can infect, changing the formula in the excel model.

The stochastic model is based on a population of 10,000 individuals. Every individual has a status as described above and has their own individual reproduction factor when infected. As mentioned above, we draw the individual reproduction factors from a gamma distribution to get the skewness effect. The infection process is modelled by randomly selecting others in the population and, if healthy, make them an infected person. To get the effect of local outbreaks the maximum number of infected is limited to 10% (1.000). The model makes several runs to see the effect of the stochastic nature of the virus. The model works with scenarios that take into account government intervention. For example if the number of infected is more than a predefined number the reproduction factor will be set to 1 due to government intervention.

We use the population of 10,000 to calculate the crown dashboard for the country by multiplying the outcomes by 1750 for the Netherlands. The confidence intervals are determined using the Central Limit Theorem.

Next steps

In this article we introduced a new Covid-19 dashboard that takes into account the stochastic nature of the virus and predicts the number of hospital and ICU beds. Policymakers can use this dashboard to plan the capacity of hospital beds and ICUs, taking into account the confidence level that they require, i.e., the level of risk they are willing to take.

To arrive at a more reliable Covid-19 dashboard, the authors recognized the following questions that need to be answered:

- Do asymptomatically sick individuals infect with the same reproduction factor as other infected people?
- What is the level of skewness of the Covid-19 infections? The confidence levels depend on the assumed level of skewness of the infections
- What is the duration of immunity and is this related to the level of sickness of an individual?

A Corona-dashboard taking into account the confidence level is a useful tool for policymakers and will even be more useful when these questions are answered so as to get smaller and more precise intervals.

For the interested reader we refer to our educational website www.crownactuaries.eu for further information and discussion.

Kees van Heugten: Kees is a qualified actuary and mathematician with more than 25 years experience at insurance companies and pension funds. His expertise lies in the field of (actuarial) modelling, reporting and risk management. In 2019 Kees graduated as a Certified Enterprise Risk Actuary (CERA). Kees is a member of the Royal Actuarial Society and the Royal Mathematical Society of the Netherlands. Kees lives in the Netherlands (Hilversum), with Wilfreda, Dara, Giulia and their two dogs Hunter and Fox.

Servaas Houben: Servaas studied econometrics in the Netherlands and worked there for the first 4 years of his career. Thereafter Servaas worked in Dublin, London, and Curacao. Besides his actuarial qualifications, Servaas completed the CFA and FRM qualifications. Servaas writes regularly for his blog (https://actuaryabroad.wordpress.com/), CFA digest, and actuarial magazines. Servaas lives in Ijsselstein with his wife Kareljan and son Max.

Rob Smit: Rob is a qualified actuary and has more than 20 years experience in the field of pricing models and reserving techniques. For five years Rob has been the owner of restaurant RoberTine in the historical center of Lisbon, Portugal. Rob was inspired by Covid-19 to understand more about the virus, even after being forced to close his restaurant temporarily due to the strict conditions towards tourism in Lisbon. Rob has run a swimming club as president for more than six years. He lives in Setúbal, Portugal with his wife Tine.