KONINKLIJKE NEDERLANDSE
AKADEMIE VAN WETENSCHAPPEN

EPIDEMICS OF INFECTIOUS DISEASES: WHAT SCIENCE SAYS

A KNAW brochure

by Christina Vandenbroucke-Grauls, Jan Vandenbroucke and Jos van der Meer, Emeritus Professors of Medical Microbiology (Amsterdam), Clinical Epidemiology (Leiden), and Internal Medicine (Nijmegen)

What do we mean by infectious diseases and epidemics?

Infectious diseases (also called communicable or transmissible diseases) are diseases that people can pass on to one another. They are caused by microorganisms ("pathogens" or "germs") such as viruses, bacteria, fungi, and parasites. Non-infectious diseases, such as most cancers, cardiovascular diseases, or Alzheimer's disease, cannot be passed from one person to another. Some types of cancer, such as cervical or liver cancer, are caused by viruses, with the infection by the virus occurring long before the disease develops.

When a disease suddenly occurs much more frequently than before, we talk of an *epidemic*, and when an epidemic occurs on a global scale, it is called a *pandemic*. "Suddenly" is a relative term here. For example, people talk about the "epidemic of lung cancer" in the 20th century, which was caused by many people smoking cigarettes, or more recently, the epidemic of obesity in young children, which is the result of overeating and reduced physical activity. This text is about epidemics of infectious diseases. Such epidemics can be short-lived – meaning that they last for days, weeks, or months – but they can sometimes last for many years.

We first discuss how infectious diseases are transmitted, and then look at the course of an epidemic of an infectious disease. After that, we discuss the various kinds of measures that can be taken to combat epidemics of infectious diseases and the social causes of new epidemics. Finally, we address the question of how knowledge is acquired about epidemics.

How are infectious diseases transmitted?

Infectious diseases can be transmitted in various different ways:

- Via blood and body fluids: through contact with infected blood (via medical procedures, tattooing, contaminated hypodermic needles, etc.), via blood transfusion, or via body fluids such as semen or vaginal fluid. The hepatitis B virus and HIV, for example, can be transmitted via all these routes.

- Via the hands. Infection via the hands may occur directly, for example when people shake hands, or indirectly, for example through food or objects that have been touched by the hands. Most viruses and bacteria that cause gastrointestinal infections are

Amsterdam, March 2021

transmitted in this way. For example, food poisoning can occur if a cook with diarrhea has not washed his hands properly before preparing the food.

- By contact through the air (aerogenous transmission). A person can transmit pathogens by coughing, sneezing, or spreading saliva when talking or singing. Someone else can then be infected by infected droplets (large or small) spreading through the air. Small droplets can cause infection over greater distances than larger ones because they remain suspended in the air longer and they “fly” further than large droplets, which fall to the ground quite quickly. Some bacteria and viruses are highly resistant to drying out and can also survive in dust. We know that in the case of open tuberculosis, for example, an opera singer can even infect someone at the back of the auditorium. The chicken pox and measles viruses can pass under a door and infect someone in another room. And if they are in dust, these viruses can also infect someone the next day. Influenza (“flu”) viruses and coronaviruses survive less well in the environment. They are mainly transmitted by people being coughed on.

- From mother to child. Some infectious diseases can be transmitted from mother to child. This can be via the placenta, but sometimes also via the mother’s breast milk. HIV, the zika virus, and the West Nile virus are examples of this.

- Via a “vector” (i.e. a disease transmitter). Some pathogens are transmitted by mosquitoes, ticks, mites, or fleas. Examples are malaria, dengue fever, and Lyme disease.

- From animals to humans (zoonoses). A vector often plays a role here as well. Examples of zoonoses are Q fever (from goats to humans) and Lyme disease (from deer via ticks to humans).

It’s important to note that many infectious diseases can be transmitted in various different ways. Intestinal infections can also be spread through the air if the patient vomits profusely and someone else is standing close by. Some respiratory tract infections can also be transmitted via the hands, for example if people who are not yet infected come into contact with a hand or object of someone who is infected and then touch their mouth, nose, or eyes with their hands.

How does an infectious disease epidemic run its course?

When a new infectious disease appears in a population, it almost always follows the same pattern. At first, it spreads slowly, from person to person, from one individual to another. This often happens unnoticed; we don’t yet know that there is a new disease. The wrong diagnosis is often made at this stage because doctors think that they are dealing with a disease that is similar to the new disease. This is the first part of the epidemic curve. As Figure 1 shows (the course of the plague epidemic in Bombay (Mumbai) in 1905), in the very first weeks the number of cases rises slowly.

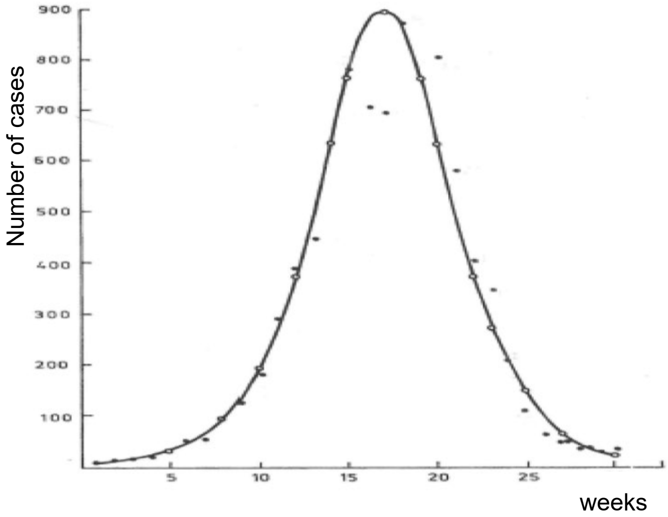

Figure 1.
Course of the 1905 plague epidemic in Bombay (Mumbai)

What happens after that stage is not easy to predict. If too few infections happen to occur (by chance), then the disease disappears again. But as soon as several people are infected and those people have become infectious themselves, the spread gets faster and faster. The number of cases increases rapidly because one person infects a number of others. This is called *exponential increase*. It can be seen in the rising line of the epidemic curve in Figure 1. At a certain point, the rate of increase slows down and the curve reaches its peak, after which the number of cases decreases.
The course of the curve is determined by the “reproduction number” (also referred to as the R). This reflects the average number of people infected by a single infected person (see also the Technical Appendix).

After the peak of the curve has been reached, the number of infections starts to decrease. There can be two reasons for this, and they can also occur simultaneously: (1) the number of susceptible people (i.e. people who can be infected) decreases, or (2) effective measures are taken to combat the epidemic. These two reasons are explained below.

1. The number of susceptible people decreases when a lot of sick people die and more and more people have already had the disease. The latter group of people are no longer susceptible: those who have survived the disease have become immune. When only very few susceptible people remain, the pathogen can no longer reach them. There are then too many people who have become immune and who now “surround” the susceptible people; in other words, there is *herd immunity*. The ultimate effect is that the epidemic disappears.

   Sometimes a disease doesn’t disappear completely but remains present. When a disease remains present for a long time with the same frequency, it is said to be *endemic*. That applies especially in the case of infectious diseases that develop slowly (for example tuberculosis or syphilis) and those transmitted by vectors (for example malaria), when there is a balance between the pathogen, the mosquito, and the population.

2. Effective measures are put in place. The simplest measure is just to go away. In the Middle Ages and the Renaissance, people who were able to (usually the rich) would leave a city where there was an epidemic of plague. During the current COVID 19 epidemic, the same thing is happening in the big American cities, with rich people moving to their second home out in the countryside.
   Most measures involve changing people's behaviour so that the epidemic "runs out of fuel" and can then no longer reach any susceptible people. For that to happen, susceptible people do not literally need to leave the city; it can also be achieved by means of other changes in behaviour. In developed countries, for example, changes in sexual behaviour and condom use have greatly reduced the HIV epidemic. Behavioural changes that can be implemented to combat the COVID-19 epidemic include "social distancing" (i.e. staying at a safe distance from one another), wearing a face mask, working from home, working behind a plastic screen, and avoiding public gatherings. These kinds of behavioural changes have the same effect as leaving a city where an epidemic is raging; they prevent transmission of the disease and the epidemic dies out.

Risk of infection

The risk of being infected by a pathogen depends on six factors:

1. the way the pathogen can reach a susceptible person;
2. the presence of one or more infectious people;
3. the quantity of pathogens that those infectious people spread around them;
4. the presence of susceptible persons, i.e. people who can still be infected;
5. the quantity of pathogens that reach those susceptible persons;
6. the quantity of pathogens needed to make someone ill (*the infectious dose*).

Let's assume, for example, that there is a pathogen that is spread by droplets (factor 1). If there are no infectious individuals in a certain space, there is no risk of infection. If there are in fact infectious people there, the risk of the disease spreading and thus of infection increases with the number of infectious individuals (factor 2) and with the quantity of pathogens spread by them (factor 3). Other relevant factors are the number of susceptible people in the vicinity of the infectious person (factor 4) and the reduction in the number of pathogens due to their becoming diluted or dying off in the surrounding area (factor 5). If, in the case of an aerogenous transmission route, there are one or more infectious individuals together with many susceptible persons in an enclosed space (i.e. factors 1 to 5 apply simultaneously), the risk of infection is extremely high. This is why gatherings of a large number of people in a concert hall, for example, quickly lead to a large number of infections. Because the dilution factor in the open air is much greater than in an enclosed space, the risk of infection in the open air is considerably lower.

Finally, there is factor 6, the infectious dose. This is the quantity of pathogens needed to make someone ill if they become infected. The infectious dose differs between pathogens. For example, taking a sip of water containing around one thousand to ten thousand *Salmonella* bacteria is enough to cause diarrhea, whereas ingesting about a million cholera bacteria – i.e. at least one hundred times as many – is necessary to develop symptoms of cholera. Someone who ingests fewer cholera bacteria doesn't become ill. Water can only contain such huge quantities of bacteria if it is heavily contaminated. The risk of that happening is mainly found in countries with a warm climate. That's why cholera only occurs in tropical countries without a good sewage system.

What measures can be taken to combat epidemics of infectious diseases?

To combat an epidemic of an infectious disease, measures need be taken to ensure that the epidemic curve rises less quickly, reaches a lower peak, and/or falls more quickly. What measures are needed depends on how most transmissions of the infection occur. How strictly the measures need to be depends on how infectious the disease is.

Diseases that are transmitted by vectors such as mosquitoes are best controlled by getting rid of the mosquitoes. Mussolini gained great popularity in 1930s Italy, for example, by having the swamps around Rome drained, thus making the city malaria-free. Blood-borne diseases also became more prevalent as surgical operations became more common in the last century. Sterilising instruments and using disposable materials meant that infection through blood could be largely prevented.

It's important for the reproduction number R – i.e. the number of new cases of a disease caused by a single infected individual (see the Technical Appendix) – to be brought below 1. If we fail to do so and each infected person infects an average of at least one other individual, then the epidemic will continue. If the R becomes smaller than 1 – even just slightly smaller – the epidemic will die out.

Isolation and quarantine: related measures, but not identical
Isolation and quarantine are related concepts, but they are clearly different from one another. Isolation is a measure that can be applied *after* it has been established that someone is ill. Quarantine is applied in the case of people who are possibly infected *before* they become ill.

Isolating infectious patients is an extremely efficient measure. It was already implemented in the Middle Ages in plague and leprosy houses, which were located outside the city. In the past, when there was an outbreak of an infectious disease (the plague, for example), the most extreme form of isolation was sometimes applied, namely the *cordon sanitaire*. That meant that the army was deployed to surround a city or a district to make sure nobody could get in or out. Isolation is still sometimes used today. At the start of the HIV epidemic in Cuba, for example, patients were isolated, with good results according to the local authorities. China is currently considering setting up large isolation centres for patients with COVID-19. In the Netherlands too, COVID-19 patients have to isolate themselves until they are no longer infectious.

The classic example of isolation as an effective measure to combat an epidemic is the way open tuberculosis was tackled in the first half of the 20th century. At that time there were no drugs or vaccines against open tuberculosis, which is spread from person to person. This form of tuberculosis is transmitted by means of aerogenous transmission, i.e. a patient coughs up bacteria and transmission occurs through coughing, singing, or speaking loudly.

By no means all cases of infection with open tuberculosis result in an infectious tuberculosis patient. A single patient with open pulmonary tuberculosis must infect an average of 20 other susceptible people for one new person to develop that form of the disease. The remaining nineteen will either overcome the bacteria on their own or will develop forms of tuberculosis that are not infectious (such as bone tuberculosis). It is therefore sufficient to ensure that patients with open pulmonary tuberculosis infect an average of fewer than twenty other people. If we persevere with that for long enough, the disease eventually disappears.

In the first half of the twentieth century, sanatoriums were set up with that aim in mind. These were hospitals located away from cities, where people suffering from open pulmonary tuberculosis were gathered together in isolation. They were well cared for, given good food, and lived in large rooms with spacious balconies. There was no medication; patients either died or recovered spontaneously, perhaps after lung surgery. The most important thing was that the patients were separated from society as long as they were infectious.

As soon as a case of open tuberculosis was diagnosed, the municipal health service (in the Netherlands the GGD) also tried to trace the sick person's contacts to see whether any of them also had open pulmonary tuberculosis. If they did, they were taken to a sanatorium too. In addition, extensive population screening was carried out among schoolchildren. Before they could be appointed, teachers and civil servants had to be able to show a "clean" X-ray of their lungs.

As a result of all these measures, tuberculosis was already on the decline before effective medication became available in the 1950s. The great advantage of medication is that even although patients are still not cured after 14 days of treatment, they are no longer infectious and can therefore return to the community. Sanatoriums thus ceased to be necessary. But even today, when a case of open tuberculosis is diagnosed, the patient's contacts are traced and examined.

As mentioned, *quarantine* is used in the case of people who are possibly infected before they become ill. The word comes from the Italian *quaranta giorni*, meaning forty days. Back in the fourteenth century, that was how long foreign ships had to lie at anchor off the port of Ragusa (Sicily) to see if anyone on board fell ill or died. Only if everyone on board remained healthy during those forty days was the entire crew allowed to come ashore.

When someone travels to the Netherlands from an area with a lot of cases of COVID-19 and he or she is not found to be infected, then he or she is asked to go into quarantine. People who may have been in contact with an infected person in the Netherlands, or who have a family member who appears to be infected, are also asked to do so. Such "home quarantine" is not currently mandatory in the Netherlands.

So, isolation and quarantine were already applied back in the Middle Ages, and they are still measures that are very effective nowadays. That became clear in Australia and New Zealand when the COVID-19 epidemic was virtually eradicated there in the course of 2020, and has since only occasionally reappeared in a small number of cases. When they arrive, people entering those countries had to – and still have to – spend 14 days under supervised quarantine; they are then only allowed out if their test result is negative. People entering China also continue to be quarantined. In China, a form of *cordon sanitaire* was applied around the city of Wuhan, and in Australia, too, cities and areas were isolated. In all these countries, these measures were accompanied by a strict countrywide lockdown. Various European countries currently have similar measures in place (almost entirely closing their borders, and with a relatively strict lockdown, with schools and shops remaining closed and with a night-time curfew). The argument for this is that the new variants of the virus may be more infectious than the original virus.

Vaccination
Vaccination is also intended to influence the course of the epidemic curve in a way that is beneficial. Vaccination reduces the number of susceptible people in a population by enabling them to build up *immunological resistance*.[i]
A major advantage of rapid, extensive vaccination campaigns is that the number of susceptible people who keep an epidemic going can be reduced much faster than if we just wait until enough people have died or have become naturally immune. (See the Technical Appendix for the relationship between herd immunity through vaccination and the reproduction number R).

Vaccination programmes can also lead to lasting herd immunity. This occurs when "newcomers" to a population (such as new-born babies) are vaccinated. If the pathogen has mutated and the vaccines used up to that point are no longer effective, repeated vaccination with a modified vaccine can induce effective protection. That's why annual flu vaccination is necessary, because the flu virus is constantly changing. In the case of measles, however, once-only vaccination is sufficient; that virus does not change (or at least it hasn't yet).

Personal hygiene
In the battle against all diseases transmitted between humans, personal hygiene is extremely important. One important measure to prevent the spread of infectious diseases is for people to wash their hands thoroughly. They also need to avoid touching their mouth, nose or eyes, because those are routes through which pathogens enter the body. It is also important to cough and sneeze into a (disposable) handkerchief or into your elbow.

What are the social causes of new epidemics?

The emergence and spread of an epidemic run parallel to social changes and developments; these are different in each century and in each society, although there are also some persistent factors.

Both in the past and today, epidemics almost always have their hardest impact on the groups with the lowest socio-economic status. The risk of transmission is highest in those groups because they necessarily live close together ("crowding") and have the worst hygiene conditions. They also often have the lowest resistance, due to undernourishment or an unhealthy diet. During the current COVID-19 epidemic, it soon became apparent in the United States, South Africa, and in fact in all low- and middle-income countries that it is actually only well-off people who are able to follow the desired "social isolation" measures.

Intensive travel and changes in morals and habits offer new opportunities for pathogens. New local conditions may also play a role. For example, abandoned car tyres in shantytowns in developing countries create a reservoir for stagnant water in which the tiger mosquitoes that transmit dengue and zika can easily breed. Another example is Legionella, waterborne bacteria that multiply in air conditioners and shower heads, for example in hotels. These bacteria can enter the lungs in tiny water droplets and cause an outbreak of a pneumonia-type illness ("Legionnaires' disease") among hotel guests. Close contact between humans and animals also entails a risk of infection. For example, large-scale livestock farming in inhabited areas of the Netherlands turned out to play a decisive role in a major epidemic of Q fever around goat farms.

When fighting an epidemic, its exponential growth needs to be slowed down as quickly as possible, so as to prevent the peak getting too high. That's why measures must be taken, right at the start, that may at first sight seem excessive. The public may find these difficult to accept.

The World Health Organization (WHO) warned back in 2018, i.e. before the COVID-19 epidemic, that it was becoming increasingly difficult to gain sufficient support for measures such as quarantine and isolation, measures which had once seemed obvious. The WHO also already warned that alongside disease epidemics, rumour epidemics could also start on social media.

Economic arguments have from way back led to problems as regards the acceptance of quarantine and isolation. In the first half of the nineteenth century, there were two theories about the infectiousness of diseases: the "miasma theory", according to which localised "vapours" emanating from the ground caused such diseases, and the "germ theory", according to which they were transmitted from person to person. If the "germ theory" was correct, quarantine and isolation made sense but if the "miasma theory" was correct, it was better to flee the unhealthy area immediately. The merchants and other progressive citizens of the time did not favour the "germ theory" – and the medical profession agreed with them. The objection to the "germ theory" was that the associated quarantine and isolation measures meant a loss of trade and thus prosperity. However, it was the "germ theory" that was correct.

How do we acquire knowledge about epidemics?

To properly understand the origin and course of epidemics, it is necessary to thoroughly study how they occur in terms of time, place, and person. The circumstances of their transmission must also be thoroughly investigated. Here is a brief overview of some key aspects.

Time, place, and person
The aspects of time, place, and person can give clues about the causes of a disease and how it spreads..

- Time: Does the infection occur in the winter or in the summer? But also: how do cases of infection progress from day to day? The time aspect also includes the incubation period and the "window of infectivity".

  The *incubation period* is the average time that elapses between the start of an infection and the occurrence of symptoms. That time is needed for the pathogens to multiply in the body into such numbers that the infected person develops symptoms. For some infectious diseases, the incubation period can be very short (one or two days for the flu) but for others it is much longer (one to two months for mononucleosis ("mono"), for example). At the start of an epidemic of a new disease, the incubation period is still unknown and has to be estimated on the basis of reports from patients. Someone says, for example, that he fell ill ten days after visiting his sick mother for an hour, while someone else says that she fell ill five days after being visited by someone who was coughing a lot. Based on such accounts, estimates are then made of the shortest, longest, and average incubation period. We need to know these so as to determine the minimum and maximum duration of quarantine.

  The *window of infectivity* is the average number of days between the point when someone becomes infectious – which can be well before symptoms appear; see above – and the eventual disappearance of the symptoms. The window of infectivity is also estimated at the start of an epidemic on the basis of accounts from patients. It is then the basis for the isolation guidelines.

- Place: Where do the cases occur? In which districts of a town, for example? In a tropical climate or in a temperate one?

For example, a new epidemic of an unknown form of encephalitis (inflammation of the brain) once occurred near a zoo in New York, where many dead crows were also found. It turned out to be caused by a virus that infected crows that came to eat in the zoo's aviaries, which housed exotic birds. Humans were then infected via mosquitoes that transmitted the disease from the crows.

Another epidemic, of red-spotted fever, occurred only in a group of adjoining apartment blocks in New York, and above all in children. Children at the same schools who lived elsewhere didn't fall ill. The pathogen turned out to be a bacterium that lived in mice in the blocks' cellars. It was transmitted to humans by mites, and especially to children who played on the ground.

– Person: Features of the person concerned such as their age or gender, but also for example their profession, can help researchers track down the routes by which a disease spreads.
  During an epidemic of MERS-CoV (a coronavirus) in Saudi Arabia in 2012, many of the patients were found to work taking care of camels. This led researchers to identify camels as a major reservoir of the MERS virus.

But not everyone who is infected develops symptoms of the illness. Infection can occur via people who are *presymptomatic*, i.e. who do not yet have any symptoms but can still become ill, or by people who remain *asymptomatic*. Such people are infected and can also become infectious, but they do not develop symptoms themselves. Whether presymptomatic and asymptomatic people are infectious and whether they are important in spreading the disease is estimated at the start of an epidemic on the basis of patient contacts. Five publications in January 2020 in the British medical journal *The Lancet* about the first patients with COVID-19 in China, showed that the virus spread very quickly, probably because patients were spreading the disease just before they developed symptoms. People who were infected but did not fall ill themselves were also found to be able to infect others (see "Further reading"). If an infectious disease spreads via infected people who do not (yet) have any symptoms, it is very difficult to combat it. Infectious diseases that are only transmitted by people with obvious symptoms, such as smallpox in the last century, are easier to combat.

Circumstances of transmission
The circumstances in which an infectious disease is spread can also help understand its origins and its course. In the case of a new respiratory tract infection, for example, we can identify whether infected people have recently been in a restaurant, a public building, or at the hairdresser's more often than people who do not have the infection. If they have, then it is clear that people in such a situation are more likely to contract the disease. Based on that information, appropriate measures can be taken as regards visiting such places. But it's important to note that this is not the same as testing close contacts (i.e. the people with whom one was in the restaurant, or other visitors to the public building).

Another example: if snacks are served at a reception and some of those attending later vomit and suffer from diarrhoea, one can ask those who became ill what snacks they ate. By comparing their answers with those from people who did not fall ill, we can determine where the "guilty" snacks came from, where they were prepared and stored, and whether there were any sick people among the staff preparing them.

In conclusion

Infectious diseases have always existed and will never disappear because pathogens have an incredible ability to adapt to changing circumstances. This adaptation is a blind process: it happens due to random changes in the genetic material of a pathogen that give the new variant a better chance of survival. These random changes happen all the time because pathogens are constantly multiplying in enormous numbers. At the same time, the circumstances to which these accidental changes can respond are also constantly changing. And completely new pathogens are also created due to genetic changes in bacteria or viruses that have not previously caused disease in humans. That's why it's of the utmost importance to continue to study epidemics carefully. Only by correctly applying tried-and-tested epidemiological principles can new epidemics be recognised as soon as possible and the right measures be taken to combat them.

See "Further reading" for more details.

We are grateful to Andra Neefjes-Borst, Sjaak Neefjes, Teun Bousema, Cornelia van Duijn, and Ron Fouchier for critically reading previous versions of the text, and to Carel Jansen for thorough final editing.

Technical Appendix: the Reproduction Number

We have referred in this text to the reproduction number as it is measured during an epidemic. Actually, there are two reproduction numbers: the "R0", which is the basic (or basal) reproduction number; and the effective reproduction number, the "R".

The R0 or basic reproduction number is a rather abstract concept. It represents the number of new cases of a disease caused by a single infected individual when the disease occurs in a non-immune population and no measures are taken. At the start of the COVID-19 epidemic, in China, such a situation probably existed for a short time. The R0 was then estimated at almost 3. To track an epidemic, scientists use the effective reproduction number R, i.e. the number measured during the epidemic. That number represents the number of cases of a disease caused by an infected person under the current circumstances, i.e. under circumstances when measures are in fact being taken and there is already partial immunity. If the R is greater than 1 the epidemic is increasing, and when it is less than 1 the epidemic is decreasing.

There is a mathematical relationship between the basic reproduction number R0 and herd immunity: to achieve herd immunity, a proportion of at least 1-1/R0 of the population must be immune. In the case of an R0 of 2.5, the required proportion is 60% of the population. In order to incorporate a safety margin, it is usually preferable to achieve a higher percentage. It's important to note that the effective reproduction number R is unrelated to immunity: if a strict lockdown causes the R to fall to 1 then, according to the formula, vaccination would no longer be necessary – but that would mean that all the lockdown measures would need to remain in place for ever.

So based on its infectiveness, COVID-19 would die out when 60-70% or more of the population are fully protected by vaccination or by having had the disease (i.e. having acquired immunity naturally). It is probably also important that people who have been vaccinated can no longer transmit (or get) the infection and that the vaccination level is "homogeneous". This means that there are no parts of the population where the vaccination rate is much lower and where the disease can therefore still rage. In the case of an extremely infectious disease like measles, a much higher percentage of vaccinated people is needed in order to eradicate the disease (>95%).

---

[i] See https://www.knaw.nl/nl/thematisch/coronavirus/de-knaw-en-het-coronavirus/vaccineren-of-niet-wat-zegt-de-wetenschap.